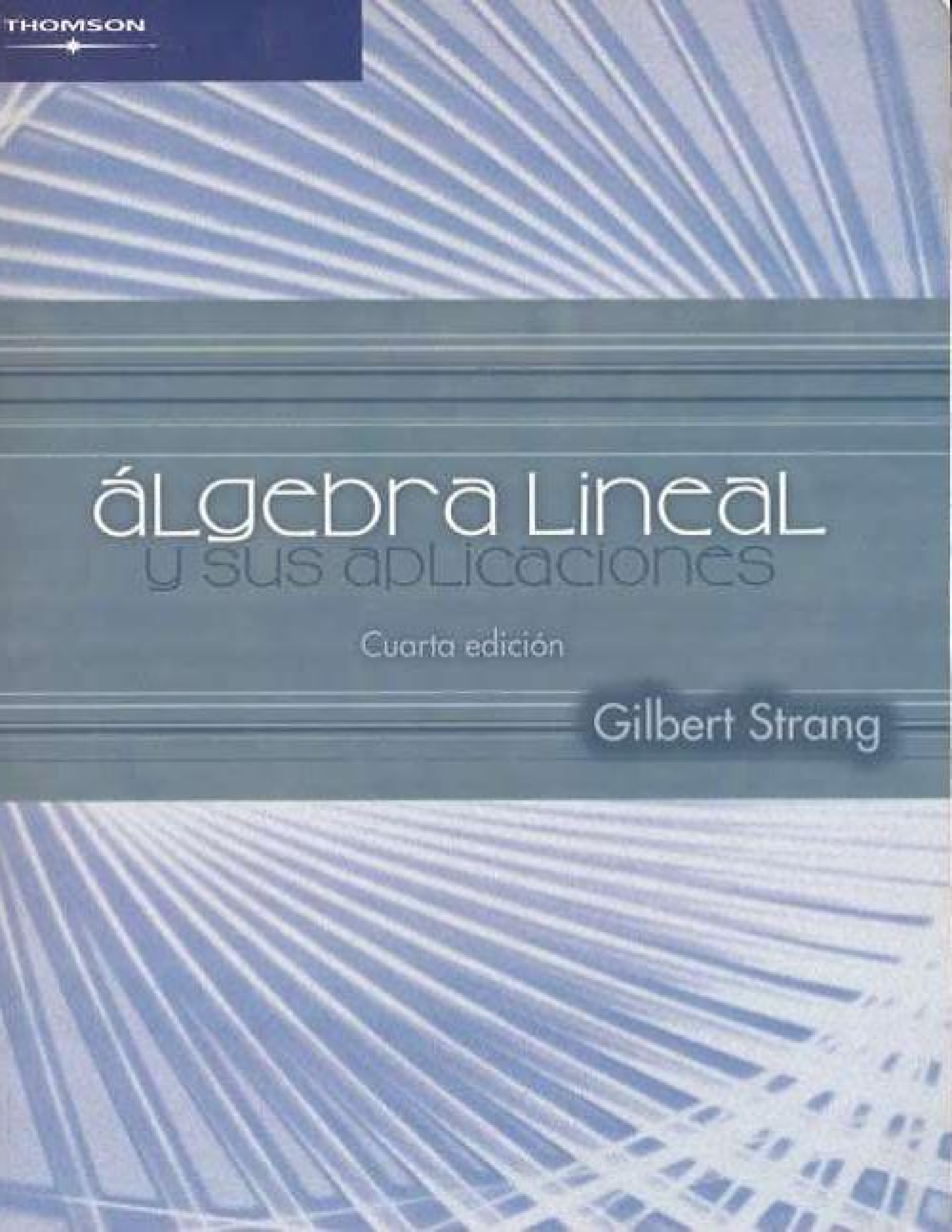

THOMSON
álgebra lineal
y sus aplicaciones
Cuarta edición
Gilbert Strang

Álgebra lineal y sus aplicaciones

cuarta edición

Gilbert Strang
Massachusetts Institute of Technology

Revisión técnica:

Edmundo Palacios Pastrana
Universidad Iberoamericana

THOMSON

Australia • Brasil • Canadá • España • Estados Unidos • México • Reino Unido • Singapur



Álgebra lineal y sus aplicaciones, 4a. Ed.
Gilbert Strang

Presidente de Thomson Learning Iberoamerica:
Javier Arellano Gutiérrez

Director editorial Iberoamérica:
José Tomás Pérez Bonilla

Gerente editorial y de producción:
Lilia Moreno Olvera

Editor de desarrollo:
Pedro de la Garza Rosales

Coordinador de preprensa:
Alejandro A. Gómez Ruiz

Coordinador de manufactura:
Israel Robles Martínez

Editora de producción:
Abril Vega Orozco

Diseño de portada:
Grupo Insigne O.T.A., S.A. de C.V.

Revisión técnica:
Edmundo Palacios Pastrana
Universidad Iberoamericana

Traducción:
Hugo Villagómez

Composición tipográfica:
Juan Castro (TROCAS)

Lectores ortotipográficos:
David García Vázquez
Felipe Martínez Martínez

Impreso en México
Printed in Mexico
1 2 3 4 09 08 07 06

Para mayor información contáctenos en:
Corporativo Santa Fe
Av. Santa Fe, núm. 505, piso 12,
Col. Cruz Manca, Santa Fe
C.P. 05349, Delegación Cuajimalpa
México, D. F.

Puede visitar nuestro sitio en
http://www.thomson.com.mx

Traducido del libro *Linear Algebra and Its Applications, 4th.*, publicado en inglés por Brooks Cole, © 2006
ISBN 0-03-010567-6
Datos para catalogación bibliográfica:
Strang, Gilbert. *Álgebra lineal y sus aplicaciones*, 4a. ed.
ISBN 970-686-609-4
Contenido:
1. Matrices y eliminación gaussiana.
2. Espacios vectoriales.
3. Ortogonalidad. 4. Determinantes.
5. Valores característicos y vectores característicos.
6. Matrices positivas definidas.
7. Cálculos con matrices. 8. Programación lineal y teoría de juegos.

División Iberoamericana

México y América Central
Thomson Learning
Corporativo Santa Fe
Av. Santa Fe, núm. 505, piso 12
Col. Cruz Manca, Santa Fe
C.P 05349, Delegación Cuajimalpa
México, D. F.
Tel. (52-55) 1500 6000
Fax (52-55) 1500 6019
editor@thomsonlearning.com.mx

El Caribe
Thomson Learning
Metro Office Park 3
Suite 201 St. 1 Lot 3
Guaynabo, Puerto Rico
Zip Code: 00968-1705
Tel. (787) 641 1112
Fax (787) 641 1119

Cono Sur
Buenos Aires, Argentina
thomson@thomsonlearning.com.ar

América del Sur
Thomson Learning
Carrera 55 núm. 67 A-05
Bogotá, Colombia
Tel. (571) 630 8212
Fax (571) 630 7999
cliente@thomsonlearning.com

España
Thomson Learning
Calle Magallanes 25
28015 Madrid
España
Tel. 34 (0) 91 446 3350
Fax 34 (0) 91 445 6218
clientes@paraninfo.es

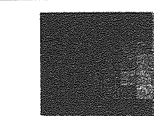


Tabla de contenido

Capítulo 1 MATRICES Y ELIMINACIÓN GAUSSIANA 1

1.1 Introducción 1
1.2 Geometría de las ecuaciones lineales 3
1.3 Un ejemplo de eliminación gaussiana 11
1.4 Notación matricial y multiplicación de matrices 19
1.5 Factores triangulares e intercambios de renglones 32
1.6 Inversas y traspuestas 45
1.7 Matrices especiales y aplicaciones 58
Ejercicios de repaso 65

Capítulo 2 ESPACIOS VECTORIALES 69

2.1 Espacios y subespacios vectoriales 69
2.2 Cómo resolver $Ax = 0$ y $Ax = b$ 77
2.3 Independencia lineal, base y dimensión 92
2.4 Los cuatro subespacios fundamentales 102
2.5 Gráficas y redes 114
2.6 Transformaciones lineales 125
Ejercicios de repaso 137

Capítulo 3 ORTOGONALIDAD 141

3.1 Vectores y subespacios ortogonales 141
3.2 Cosenos y proyecciones sobre rectas 152
3.3 Proyecciones y mínimos cuadrados 160
3.4 Bases ortogonales y Gram–Schmidt 174
3.5 La transformada discreta de Fourier 188
Ejercicios de repaso 198

Capítulo 4 DETERMINANTES 201

4.1 Introducción 201
4.2 Propiedades del determinante 203
4.3 Fórmulas para el determinante 210
4.4 Aplicaciones de los determinantes 220
Ejercicios de repaso 230

Capítulo 5 VALORES CARACTERÍSTICOS Y VECTORES CARACTERÍSTICOS 233

5.1 Introducción 233
5.2 Diagonalización de una matriz 245
5.3 Ecuaciones en diferencias y potencias A^k 254
5.4 Ecuaciones diferenciales y e^{At} 266
5.5 Matrices complejas 280
5.6 Transformaciones de semejanza 293
Ejercicios de repaso 307

Capítulo 6 MATRICES POSITIVAS DEFINIDAS 311

6.1 Mínimos, máximos y puntos silla 311
6.2 Pruebas para comprobar si una matriz es positiva definida 318
6.3 Descomposición del valor singular 331
6.4 Principios minimales 339
6.5 El método del elemento finito 346

Capítulo 7 CÁLCULOS CON MATRICES 351

7.1 Introducción 351
7.2 Norma de una matriz y número de condición 352
7.3 Cálculo de valores característicos 359
7.4 Métodos iterativos para $Ax = b$ 367

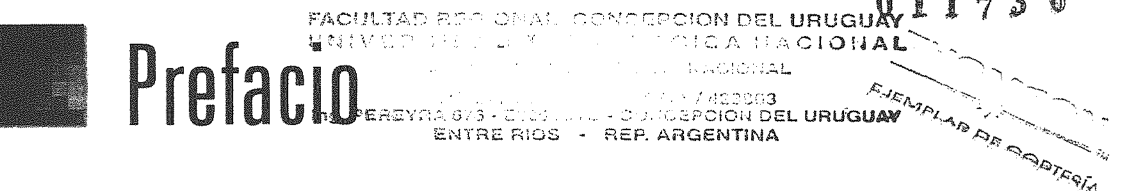
Capítulo 8 PROGRAMACIÓN LINEAL Y TEORÍA DE JUEGOS 377

8.1 Desigualdades lineales 377
8.2 El método simplex 382
8.3 El problema dual 392
8.4 Modelos de redes 401
8.5 Teoría de juegos 408

Apéndice A INTERSECCIÓN, SUMA Y PRODUCTO DE ESPACIOS 415

Apéndice B LA FORMA DE JORDAN 422

Soluciones a ejercicios seleccionados 428
Factorizaciones matriciales 474
Glosario: Un diccionario de álgebra lineal 476
Códigos de enseñanza MATLAB 481
Índice 482
Álgebra lineal en pocas palabras 488



Prefacio

La revisión de este libro de texto ha sido un desafío especial por una razón bastante agradable. Mucha gente ha leído este libro, enseñado a partir de él e inclusive lo ha querido. Quizá el espíritu del libro no cambie jamás. Este texto fue escrito como ayuda para que nuestra enseñanza del álgebra lineal mantenga la importancia crucial de este tema, que sigue creciendo.

Ciertamente, un paso era posible y aconsejable: *añadir nuevos problemas*. Tantos años de enseñanza requirieron cientos de reactivos de examen nuevos (especialmente con interrogantes que impliquen el uso de la red). Considero que el lector aprobará la amplia gama de problemas. Los reactivos siguen siendo una mezcla de *explicación* y *cálculo*: los dos métodos complementarios para aprender este hermoso tema.

Personalmente considero que mucha más gente necesita álgebra lineal que cálculo. ¡Isaac Newton podría no estar de acuerdo! Sin embargo, él no está enseñando matemáticas en el siglo XXI (y quizá no fue un gran profesor, aunque le otorgaremos el beneficio de la duda). Ciertamente, las leyes de la Física se expresan bien mediante ecuaciones diferenciales. Newton requirió del cálculo, lo cual está bien. Pero el alcance de la ciencia, la ingeniería y la administración (así como de la vida) actualmente es mucho más grande, y el álgebra lineal se ha desplazado a un sitio nodal.

Podría decir algo más, ya que muchas universidades aún no ajustan el equilibrio hacia el álgebra lineal. Al trabajar con líneas curvas y superficies curvas, el primer paso siempre es *linealizar*. Sustituir la curva por su recta tangente, ajustar la superficie por un plano y entonces el problema se vuelve lineal. El poder de este tema se hace evidente cuando se tienen 10 variables, o 1000, en vez de dos.

Quizá el lector piense que estoy exagerando cuando uso la palabra "hermoso" para un curso básico de matemáticas. En absoluto. Este curso empieza con dos vectores v y w que apuntan en direcciones distintas. El paso clave es *tomar sus combinaciones lineales*. Se multiplica para obtener $3v$ y $4w$, y se suma para obtener una combinación particular $3v + 4w$. Este nuevo vector está en el *mismo plano* que v y w. Cuando se toman todas las combinaciones, se está llenando todo el plano. Si v y w se dibujan en esta página, sus combinaciones $cv + dw$ llenan la página (y más allá), pero *no salen* de la página.

En el lenguaje de las ecuaciones lineales, $cv + dw = b$ puede resolverse exactamente cuando el vector b está en el mismo plano que v y w.

Matrices

Se avanzará un poco más para convertir combinaciones de vectores tridimensionales al álgebra lineal. Si los vectores son $v = (1, 2, 3)$ y $w = (1, 3, 4)$, se escriben en una **matriz columna**:

$$\textbf{matriz} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 4 \end{bmatrix}.$$

Para encontrar combinaciones de estas columnas, la matriz se **"multiplica"** por un vector (c, d):

$$\textbf{Combinaciones lineales } cv + dw \qquad \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = c \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + d \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}.$$

Estas combinaciones llenan un *espacio vectorial*, denominado **espacio columna** de la matriz. (Para estas dos columnas, dicho espacio es un plano.) Para decidir si $b = (2, 5, 7)$ está en ese plano, se cuenta con tres componentes para lograrlo. Así, hay que resolver tres ecuaciones:

$$\begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 7 \end{bmatrix} \quad \text{significa} \quad \begin{aligned} c + d &= 2 \\ 2c + 3d &= 5 \\ 3c + 4d &= 7. \end{aligned}$$

Se deja que el lector las resuelva. El vector $b = (2, 5, 7)$ no está en el plano de v y w. Si el 7 se cambia por cualquier otro número, entonces b no está en el plano; de hecho, *no* es ninguna combinación de v y w, por lo que las tres ecuaciones no tienen solución.

Ahora es posible describir la primera parte del libro, sobre ecuaciones lineales $Ax = b$. La matriz A tiene n columnas y m renglones. *El álgebra lineal se desplaza de manera continua hacia n vectores en el espacio m-dimensional.* Siguen buscándose combinaciones de las columnas (en el espacio columna). Siguen obteniéndose m ecuaciones para producir b (una por cada renglón). Estas ecuaciones pueden o no tener una solución. Siempre tienen una solución por mínimos cuadrados.

La interacción de las columnas y los renglones constituye el núcleo del álgebra lineal. No es totalmente fácil comprenderlo, aunque tampoco es tan difícil. A continuación se enumeran cuatro de los conceptos más importantes:

1. El **espacio columna** (todas las combinaciones de las columnas).
2. El **espacio renglón** (todas las combinaciones de los renglones).
3. El **rango** (el número de columnas independientes) (o renglones).
4. **Eliminación** (la forma idónea para encontrar el rango de una matriz).

Y aquí me detengo para permitirle iniciar el curso.

Páginas web

Quizá sea útil mencionar las páginas de la red conectadas con este libro. Recibimos muchos mensajes con sugerencias y palabras de aliento, por lo que espero que el lector utilice todo con libertad. Puede acceder directamente a *http://web.mit.edu/18.06*, que se actualiza constantemente para el curso que se imparte cada semestre. Álgebra lineal también está en el sitio del MIT OpenCourseWare *http://ocw.mit.edu*, donde 18.06 se volvió excepcional al incluir videos de las conferencias (que, por supuesto, usted no tiene que ver...). A continuación se menciona una parte del material disponible en la red:

1. Programa de conferencias y tareas y exámenes actuales con soluciones.
2. Los objetivos del curso, así como preguntas conceptuales.
3. Demos interactivos Java (ahora ya se cuenta con audio para los valores característicos).
4. Códigos de enseñanza del álgebra lineal y problemas MATLAB.
5. Videos de todo el curso (tal y como se enseña en un aula real).

La página del curso se ha convertido en un vínculo valioso para la clase y un recurso para los estudiantes. Estoy bastante optimista sobre el potencial de las gráficas sonoras. El ancho de banda para la voz en off es bajo, y FlashPlayer está disponible de manera gratuita. Esto ofrece un *repaso rápido* (con experimentos activos), y es posible bajar todas las conferencias. Espero que los profesores y estudiantes de todo el mundo encuentren útiles estas páginas web. Mi objetivo es hacer este libro lo más útil posible con todo el material del curso que puedo proporcionar.

Nota importante: La administración de estas páginas Web y de otras mencionadas dentro del libro no está a cargo de Thomson Learning Iberoamérica, por lo que la editorial no es responsable de las modificaciones en el contenido y los cambios en las políticas y formas de acceso que pudieran ocurrir. Le recomendamos visitar frecuentemente estos sitios a fin de estar al tanto de cualquier actualización.

Complementos

Este libro cuenta con complementos para el profesor, los cuales están en inglés y sólo se proporcionan a los docentes que adopten la presente obra como texto para sus cursos. Para mayor información, favor de comunicarse con las oficinas de nuestros representantes o a los siguientes correos electrónicos:

Thomson México y Centroamérica clientes@thomsonlearning.com.mx.
Thomson América del Sur cliente@thomsonlearning.com
Thomson Caribe amy.reyes@thomsonlearning.com

Estructura del curso

Los dos problemas fundamentales son $Ax = b$ y $Ax = \lambda x$ para matrices cuadradas A. El primer problema $Ax = b$ tiene una solución cuando las *columnas de A son independientes*. El segundo problema $Ax = \lambda x$ es para *vectores característicos independientes*. Una parte crucial de este curso es aprender el significado de "independencia".

Considero que la mayoría de nosotros aprendemos primero a partir de ejemplos. Puede ver que

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \end{bmatrix} \qquad \textbf{\textit{no} tiene columnas independientes.}$$

La columna 1 más la columna 2 es igual a la columna 3. Un teorema maravilloso del álgebra establece que los tres renglones tampoco son independientes. El tercer renglón debe estar en el mismo plano que los dos primeros renglones. Con alguna combinación de los renglones 1 y 2 se obtiene el renglón 3. Quizá el lector pueda encontrar rápidamente esta combinación (yo no pude). Al final tuve que usar eliminación para descubrir que la combinación correcta utiliza 2 veces el renglón 2, menos el renglón 1.

La eliminación es la forma simple y natural para entender una matriz al producir bastantes elementos iguales a cero. Por tanto, el curso empieza aquí. ¡Pero no se quede demasiado aquí! El lector debe proceder de combinaciones de los renglones a independencia de los renglones a la "dimensión del espacio renglón". Este es el objetivo clave, abordar todos los espacios de los vectores: el ***espacio renglón***, el ***espacio columna*** y el ***espacio nulo***.

Otro objetivo es comprender la manera en que *actúa* la matriz. Cuando A se multiplica por x se obtiene el nuevo vector Ax. Todo el espacio de vectores se mueve; es "transformado" por A. Transformaciones especiales se obtienen de matrices particulares, y aquéllas son las primeras piedras del álgebra lineal: matrices diagonales, matrices ortogonales, matrices triangulares, matrices simétricas.

Los valores característicos de estas matrices también son importantes. Considero que las matrices de 2 por 2 constituyen ejemplos contundentes de la información que pueden proporcionar los valores característicos λ. Las secciones 5.1 y 5.2 ameritan una lectura cuidadosa para ver la manera en que $Ax = \lambda x$ es de utilidad. Ahí se presenta un caso en que matrices pequeñas permiten la obtención de muchísimo conocimiento.

En forma global, la belleza del álgebra lineal puede apreciarse de varias maneras:

1. **Visualización.** Combinación de vectores. Espacios de vectores. Rotación, reflexión y proyección de vectores. Vectores perpendiculares. Cuatro subespacios fundamentales.

2. **Abstracción.** Independencia de vectores. Base y dimensión de un espacio vectorial. Transformaciones lineales. Descomposición del valor singular y la mejor base.

3. **Cálculo.** Eliminación para producir elementos cero. Gram-Schmidt para producir vectores ortogonales. Valores característicos para resolver ecuaciones diferenciales y en diferencias.

4. **Aplicaciones.** Solución por mínimos cuadrados cuando $Ax = b$ tiene demasiadas ecuaciones. Ecuaciones en diferencias que aproximan ecuaciones diferenciales. Matrices de probabilidad de Markov (¡la base para Google!). Vectores característicos ortogonales como ejes principales (y más . . .).

Para continuar con estas aplicaciones, se mencionan los libros publicados por Wellesley-Cambridge Press. Todos aparentan ser de álgebra, aplicados al procesamiento de señales, a ecuaciones diferenciales parciales y a cálculos científicos (e inclusive GPS: Sistema de Posicionamiento Global). Si el lector consulta la página *http://www.wellesleycambride.com*, verá parte de la razón por la que el álgebra lineal es tan utilizada.

Después de este prefacio, el libro hablará por sí mismo. De inmediato observará el espíritu. El énfasis se pone en la comprensión: *intento explicar, más que deducir*. Este es un libro sobre matemáticas verdaderas, no un ejercicio interminable. En clase, constantemente trabajo con ejemplos para enseñar lo que necesitan los estudiantes.

Agradecimientos

Disfruté la redacción de este libro, y ciertamente espero que el lector disfrute leyéndolo. Mucho de este placer proviene por haber trabajado con amigos. Recibí una ayuda maravillosa de Brett Coonley, Cordula Robinson y Erin Maneri, quienes crearon los archivos LaTeX y trazaron todas las figuras. Sin el apoyo constante de Brett nunca hubiera podido terminar esta nueva edición.

Steven Lee y Cleve Moler me proporcionaron ayuda previa con los códigos de enseñanza. Ellos siguieron los pasos descritos en el libro: MATLAB, Maple y Mathematica son más rápidos para matrices grandes. Todos pueden utilizarse (*de manera opcional*) en este curso. Hubiera podido añadir "Factorización" a la lista anterior, como una quinta avenida para la comprensión de las matrices:

[L,U,P]	= lu(A)	para ecuaciones lineales
[Q,R]	= qr(A)	para hacer ortogonales a las columnas
[S,E]	= eig(A)	para encontrar vectores característicos y valores característicos.

Al dar las gracias, nunca me olvido de la primera dedicatoria de este libro, hace años. Aquélla fue una oportunidad especial para agradecer a mis padres por haberme proporcionado tantos presentes generosos. Su ejemplo es una inspiración en mi vida.

Y también agradezco al lector, esperando que le agrade este libro.

Gilbert Strang

Capítulo 1

Matrices y eliminación gaussiana

1.1 INTRODUCCIÓN

Este libro empieza con el problema central del álgebra lineal: *la solución de ecuaciones lineales*. El caso más importante, y el más sencillo, es cuando el número de incógnitas es igual al número de ecuaciones. Se tienen ***n* ecuaciones en *n* incógnitas**, empezando con $n = 2$:

$$\textbf{Dos ecuaciones} \quad 1x + 2y = 3 \qquad \textbf{Dos incógnitas} \quad 4x + 5y = 6. \tag{1}$$

Las incógnitas son x y y. Para resolver estas ecuaciones requiero, describir dos métodos el de *eliminación* y el de *determinantes*. Ciertamente, x y y están determinadas por los números 1, 2, 3, 4, 5, 6. La cuestión es cómo utilizar estos seis números para resolver el sistema.

1. Eliminación De la segunda ecuación, réstese 4 veces la primera ecuación. Así se elimina x de la segunda ecuación, y queda una ecuación para y:

$$\textbf{(ecuación 2)} - \textbf{4(ecuación 1)} \qquad -3y = -6. \tag{2}$$

De inmediato se sabe que $y = 2$. Luego, x se conoce a partir de la primera ecuación, $1x + 2y = 3$:

$$\textbf{Con sustitución hacia atrás} \qquad 1x + 2(2) = 3 \qquad \text{se obtiene} \qquad x = -1. \tag{3}$$

Procediendo cuidadosamente, se comprueba que x y y también resuelven la segunda ecuación. Esto debe funcionar, como es el caso: 4 veces ($x = -1$) más 5 veces ($y = 2$) es igual a 6.

2. Determinantes La solución $y = 2$ depende completamente de los seis números en las ecuaciones. Debe haber una fórmula para y (y también para x). Se trata de una "razón de determinantes", que espero, el lector me permita escribir directamente:

$$y = \frac{\begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}} = \frac{1 \cdot 6 - 3 \cdot 4}{1 \cdot 5 - 2 \cdot 4} = \frac{-6}{-3} = 2. \tag{4}$$

Lo anterior puede parecer algo misterioso, a menos que el lector ya conozca algo sobre determinantes de 2 por 2. Estos determinantes proporcionan la misma respuesta $y = 2$, proveniente de la misma razón de –6 a –3. Si nos quedamos con los determinantes (lo cual no pensamos hacer), hay una fórmula semejante para calcular la otra incógnita, x:

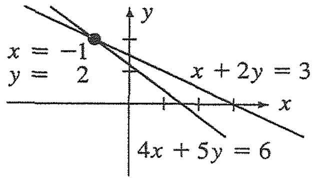
$$x = \frac{\begin{vmatrix} 3 & 2 \\ 6 & 5 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}} = \frac{3 \cdot 5 - 2 \cdot 6}{1 \cdot 5 - 2 \cdot 4} = \frac{3}{-3} = -1. \qquad (5)$$

A continuación se compararán ambos métodos, pensando en futuros problemas reales en los que n es mucho más grande ($n = 1000$ es un tamaño bastante moderado en cálculos científicos). Lo cierto es que el uso directo de la fórmula de los determinantes para 1000 ecuaciones puede ser un desastre total, ya que el millón de números a la izquierda se utilizaría correcta pero ineficazmente. Esta fórmula se encontrará en el capítulo 4 (regla de Cramer), aunque en el capítulo 1 se presenta un método aceptable para resolver 1000 ecuaciones.

Este método aceptable es la ***eliminación gaussiana***. Se trata del algoritmo que suele aplicarse de manera constante para resolver grandes sistemas de ecuaciones. A partir de los ejemplos en un libro de texto ($n = 3$ se aproxima al límite superior de la paciencia del autor y del lector), quizá el lector no puede apreciar mucha diferencia. En las ecuaciones (2) y (4) se siguieron esencialmente los mismos pasos para encontrar $y = 2$. Ciertamente, x se conoció más rápido por la sustitución hacia atrás en la ecuación (3) que la razón en (5). Para n más grande, no hay caso. Gana la eliminación (e incluso este método es el mejor para calcular determinantes).

La idea de eliminación es engañosamente simple: el lector la dominará luego de unos cuantos ejemplos. Constituye la base de la mitad de este libro, simplificando una matriz de modo que sea posible comprenderla. Junto con la mecánica del algoritmo, en este capítulo es necesario explicar cuatro aspectos más profundos. Éstos son:

1. Las ecuaciones lineales llevan a la ***geometría de planos***. No es fácil visualizar un plano nueve-dimensional en un espacio de diez dimensiones. Es más difícil ver diez de estos planos, que se cortan en la solución de diez ecuaciones, aunque de alguna manera esto es casi posible. Nuestro ejemplo tiene dos rectas en la figura 1.1, que se encuentran en el punto $(x, y) = (-1, 2)$. El álgebra lineal mueve esta imagen hacia diez dimensiones, donde la intuición debe imaginar la geometría (y la obtiene correctamente).
2. Pasamos a la ***notación matricial*** al escribir las n incógnitas como un vector x y las n ecuaciones como $Ax = b$. Multiplicamos A por "matrices de eliminación" con la finalidad de obtener una matriz triangular superior U. Con lo anterior, A se factoriza en L *veces* U, donde L es triangular inferior.



Una solución $(x, y) = (-1, 2)$

Paralelas: no hay solución

Recta completa de soluciones

Figura 1.1 El ejemplo tiene una solución. Los casos singulares no tienen solución, o tienen demasiadas soluciones.

A continuación se escribirán A y sus factores para nuestro ejemplo, y se explicarán en su debido momento:

$$\textbf{Factorización} \quad A = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix} = L \textit{ veces } U. \qquad (6)$$

Primero es necesario presentar las matrices, los vectores y las reglas de la multiplicación. Toda matriz tiene una ***traspuesta*** A^T. Esta matriz, tiene una ***inversa*** A^{-1}.

3. En la mayor parte de los casos, la eliminación se realiza sin dificultades. La matriz tiene una inversa, y el sistema $Ax = b$ tiene una solución. En casos excepcionales el método *falla*; ya sea que las ecuaciones se escribieron en orden equivocado, lo cual se arregla fácilmente al intercambiarlas, o las ecuaciones no tienen una solución única.

   El ***caso singular*** aparece si 8 se sustituye por 5 en nuestro ejemplo:

$$\begin{array}{ll} \textbf{Caso singular} & 1x + 2y = 3 \\ \textbf{Dos rectas paralelas} & 4x + 8y = 6. \end{array} \qquad (7)$$

   La eliminación resta simplemente 4 veces la primera ecuación de la segunda. Sin embargo, ¡observe el resultado!

$$\textbf{(ecuación 2) – 4(ecuación 1)} \qquad 0 = -6.$$

   Este caso singular ***no tiene solución***. Otros casos singulares tienen una ***infinidad de soluciones***. (Cambie 6 a 12 en el ejemplo, y la eliminación producirá $0 = 0$. Así, y, puede asumir *cualquier valor*.) Cuando la eliminación falla, se quiere encontrar toda solución posible.

4. Se requiere una estimación aproximada del ***número de pasos de eliminación*** necesarios para resolver un sistema de tamaño n. El costo de cómputo a menudo determina la precisión del modelo. Cien ecuaciones requieren alrededor de 300 000 pasos (multiplicaciones y restas). La computadora es capaz de hacer estos pasos rápidamente, pero no es así para el caso de varios billones de pasos. Y después de un millón de pasos, el error por redondeo puede ser significativo. (Algunos problemas son sensibles; otros no.) Sin entrar en todos los detalles, pretendemos considerar grandes sistemas que se presentan en la práctica, así como la manera en que se resuelven realmente.

El resultado final de este capítulo, es un algoritmo de eliminación que es casi lo más eficaz posible. Se trata del algoritmo que suele usarse en una numerosa variedad de aplicaciones. Y al mismo tiempo, comprenderlo en términos de *matrices* —la matriz de coeficientes A, las matrices E y P para la eliminación e intercambio de renglones, respectivamente, y los factores finales L y U— es un fundamento esencial de la teoría. Espero que el lector disfrute este libro y su curso.

1.2 GEOMETRÍA DE LAS ECUACIONES LINEALES

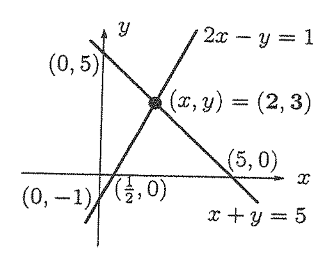
La forma de comprender este tema es mediante un ejemplo. Se empieza con dos ecuaciones extremadamente simples, reconociendo que el lector puede resolverlas sin necesidad de llevar un curso de álgebra lineal. No obstante, espero que le dé una oportunidad a Gauss:

$$\begin{aligned} 2x - y &= 1 \\ x + y &= 5. \end{aligned}$$

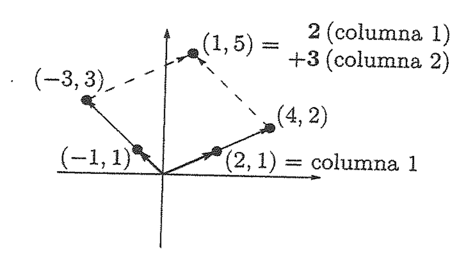
Este sistema puede abordarse *por renglones* o *por columnas*. Queremos abordarlo en ambas formas.

El primer método se centra en las ecuaciones por separado (los ***renglones***). Es el más conocido, y en dos dimensiones se hace rápidamente. La ecuación $2x - y = 1$ se representa por una *línea recta* en el plano *x-y*. La recta pasa por los puntos $x = 1$, $y = 1$ y $x = \frac{1}{2}$, $y = 0$ (y también por (2, 3) y todos los puntos intermedios). La segunda ecuación, $x + y = 5$, produce una segunda recta (véase la figura 1.2*a*). Su pendiente es $dy/dx = -1$ y corta a la primera recta en la solución.

El punto de intersección pertenece a ambas rectas. Se trata de la única solución de las dos ecuaciones. El punto $x = 2$ y $y = 3$ se encontrará pronto por "eliminación".



a) Las rectas se cortan en $x = 2$, $y = 3$



b) Las columnas se combinan con 2 y 3

Figura 1.2 Representación por renglón (dos rectas) y representación por columna (se combinan columnas).

El segundo método considera las ***columnas*** del sistema lineal. Las dos ecuaciones por separado en realidad son ***una ecuación vectorial***:

$$\textbf{Forma de columna} \qquad x\begin{bmatrix}2\\1\end{bmatrix} + y\begin{bmatrix}-1\\1\end{bmatrix} = \begin{bmatrix}1\\5\end{bmatrix}.$$

El problema consiste en ***encontrar la combinación de los vectores columna en el miembro izquierdo que produce el vector del miembro derecho***. Los vectores (2, 1) y (−1, 1) se representan con las líneas gruesas en la figura 1.2b. Las incógnitas son los números *x* y *y* que multiplican a los vectores columna. Toda la idea puede verse en esa figura, donde 2 veces la columna 1 se suma a 3 veces la columna 2. Geométricamente, así se obtiene un famoso paralelogramo. Algebraicamente, se obtiene el vector correcto (1, 5), en el miembro derecho de las ecuaciones. La representación por columnas confirma que $x = 2$ y $y = 3$.

Puede dedicarse más tiempo a este ejemplo, aunque se prefiere pasar al caso en que $n = 3$. Tres ecuaciones siguen siendo manipulables, y presentan mucha mayor variedad:

$$\textbf{Tres planos} \qquad \begin{aligned} 2u + v + w &= 5 \\ 4u - 6v \phantom{{}+ w} &= -2 \\ -2u + 7v + 2w &= 9. \end{aligned} \qquad (1)$$

De nuevo, es posible estudiar los renglones o las columnas, de modo que se empieza con los renglones. Cada ecuación describe un ***plano*** en tres dimensiones. El primer plano es $2u + v + w = 5$, y se muestra en la figura 1.3. Contiene a los puntos $(\frac{5}{2}, 0, 0)$ y (0, 5, 0) y (0, 0, 5). Es determinado por cualesquiera tres de sus puntos, suponiendo que no son colineales.

Al cambiar 5 a 10, el plano $2u + v + w = 10$ *debe ser paralelo al anterior*. Contiene a (5, 0, 0) y (0, 10, 0), y (0, 0, 10) que está dos veces más lejos del origen, que es el

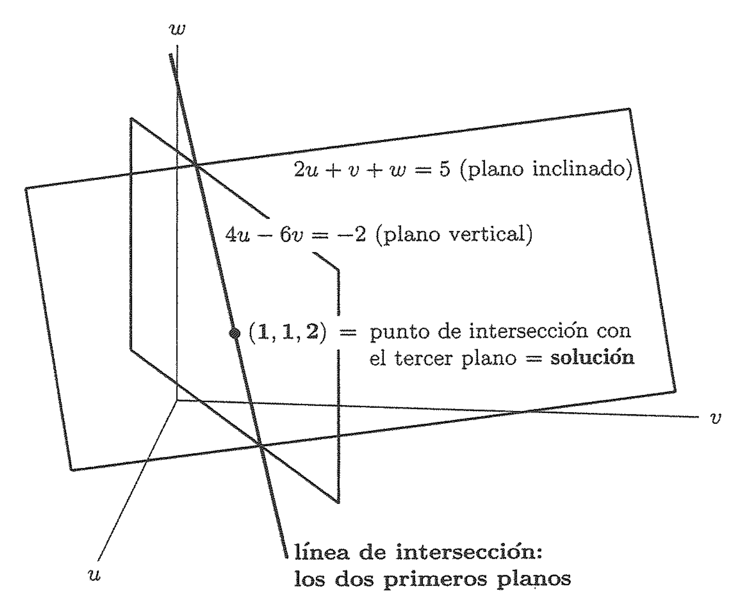


Figura 1.3 La representación por renglón: tres planos que se cortan, provenientes de tres ecuaciones lineales.

punto central $u = 0$, $v = 0$, $w = 0$. Al cambiar el miembro derecho, el plano paralelo se mueve a sí mismo, y el plano $2u + v + w = 0$ pasa por el origen.

El segundo plano es $4u - 6v = -2$. Se trazó verticalmente porque *w* puede asumir cualquier valor. El coeficiente de *w* es cero, aunque sigue siendo un plano en el espacio tridimensional. (La ecuación $4u = 3$, o incluso el caso extremo $u = 0$, sigue describiendo un plano.) En la figura se muestra la intersección del segundo plano con el primero. Esta intersección es una recta. *En tres dimensiones, una recta requiere dos ecuaciones*; en *n* dimensiones requiere $n - 1$.

Finalmente, el tercer plano corta a la recta en un punto. El plano (no está dibujado) representa la tercera ecuación, $-2u + 7v + 2w = 9$, y corta a la recta en $u = 1$, $v = 1$, $w = 2$. Este punto de intersección triple (1, 1, 2) resuelve el sistema lineal.

¿Cómo se extiende hasta *n* dimensiones esta representación por renglones? Las *n* ecuaciones contienen *n* incógnitas. La primera ecuación sigue determinando un "plano", que ya no es un plano bidimensional en el espacio tridimensional; de alguna manera, su "dimensión" es $n - 1$. Debe ser plano y extremadamente delgado en un espacio *n*-dimensional, aunque nos parezca sólido.

Si el tiempo es la cuarta dimensión, entonces el plano $t = 0$ corta al espacio tetradimensional y produce el universo tridimensional en que vivimos (o más bien, el universo como era en $t = 0$). Otro plano es $z = 0$, que también es tridimensional; se trata del plano normal *x-y* que se considera todo el tiempo. ¡Estos espacios tridimensionales se cortan! Comparten el plano normal *x-y* en $t = 0$. Si se desciende a dos dimensiones, el siguiente plano deja una recta. Por último, un cuarto plano deja un solo punto. Se trata del punto de intersección de 4 planos en 4 dimensiones, y resuelve las 4 ecuaciones subyacentes.

Si continuamos con este ejemplo proveniente de la relatividad, estaremos en problemas. La cuestión es que el álgebra lineal es capaz de operar con cualquier número de ecuaciones. La primera ecuación produce un plano $(n - 1)$ dimensional en *n* dimensiones. El

segundo plano lo corta (esperamos) un conjunto más pequeño de "dimensión $n - 2$". Suponiendo que todo va bien, todo plano nuevo (toda ecuación nueva) reduce la dimensión en una unidad. Al final, cuando se hayan tomado en cuenta todos los n planos, la dimensión de la intersección es cero. Se trata de *un punto*, que pertenece a todos los planos, y sus coordenadas satisfacen a todas las n ecuaciones. ¡Ésta, es la solución!

Vectores columna y combinaciones lineales

Ahora volvemos a las columnas. Esta vez la ecuación vectorial (la misma ecuación que (1)) es

$$\textbf{Forma de columna} \qquad u\begin{bmatrix}2\\4\\-2\end{bmatrix} + v\begin{bmatrix}1\\-6\\7\end{bmatrix} + w\begin{bmatrix}1\\0\\2\end{bmatrix} = \begin{bmatrix}5\\-2\\9\end{bmatrix} = b. \qquad (2)$$

Estos *son vectores columna tridimensionales*. ***El vector b se identifica con el punto cuyas coordenadas son 5, −2, 9***. Todo punto en el espacio tridimensional se hace corresponder con un vector, y viceversa. Ésta, era la idea de Descartes, quien transformó la geometría en álgebra al trabajar con las coordenadas del punto. Es posible escribir el vector en una columna, o sus componentes pueden enumerarse como $b = (5, -2, 9)$, o incluso puede representarse geométricamente mediante una flecha a partir de su origen. Pueden elegirse *la flecha*, o *el punto* o *los tres números*. En seis dimensiones, quizá es más conveniente elegir los seis números.

Cuando los componentes se enumeran horizontalmente, suele utilizarse paréntesis y comas, y cuando el vector columna se indica verticalmente se usan llaves (sin comas). Lo que realmente importa es la ***suma de vectores*** y la ***multiplicación por un escalar*** (un número). En la figura 1.4a se muestra una suma vectorial, componente por componente:

$$\textbf{Suma vectorial} \qquad \begin{bmatrix}5\\0\\0\end{bmatrix} + \begin{bmatrix}0\\-2\\0\end{bmatrix} + \begin{bmatrix}0\\0\\9\end{bmatrix} = \begin{bmatrix}5\\-2\\9\end{bmatrix}.$$

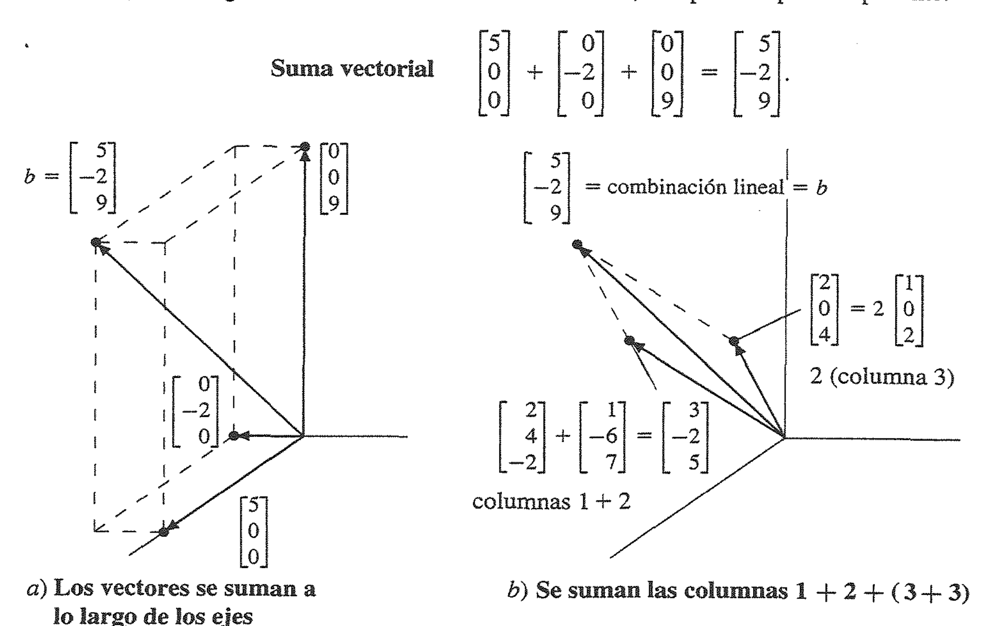


Figura 1.4 La representación por columna: la combinación lineal de las columnas es igual a b.

En la figura de la derecha hay una multiplicación por 2 (y si el vector hubiese sido multiplicado por –2, entonces el vector hubiera invertido su dirección):

$$\textbf{Multiplicación por escalares} \qquad 2\begin{bmatrix}1\\0\\2\end{bmatrix} = \begin{bmatrix}2\\0\\4\end{bmatrix}, \qquad -2\begin{bmatrix}1\\0\\2\end{bmatrix} = \begin{bmatrix}-2\\0\\-4\end{bmatrix}.$$

También en la figura de la derecha se observa una de las ideas centrales del álgebra. Utiliza *las dos* operaciones básicas: los vectores *se multiplican por números y luego se suman*. El resultado se denomina ***combinación lineal***, y esta combinación resuelve nuestra ecuación:

$$\textbf{Combinación lineal} \qquad \mathbf{1}\begin{bmatrix}2\\4\\-2\end{bmatrix} + \mathbf{1}\begin{bmatrix}1\\-6\\7\end{bmatrix} + \mathbf{2}\begin{bmatrix}1\\0\\2\end{bmatrix} = \begin{bmatrix}5\\-2\\9\end{bmatrix}.$$

La ecuación (2) requirió multiplicadores u, v, w, que producen el miembro derecho b. Estos números son $u = \mathbf{1}$, $v = \mathbf{1}$, $w = \mathbf{2}$. Y proporcionan la combinación correcta de las columnas. También proporcionaron el punto (1, 1, 2) en la representación por renglón (donde se cortan los tres planos).

Nuestro verdadero objetivo es ir más allá de dos o tres dimensiones, hasta n dimensiones. Con n ecuaciones en n incógnitas, en la representación por renglón hay n planos. En la representación por columna hay n vectores, más un vector b en el miembro derecho. La ecuación pide ***una combinación lineal de las n columnas que sea igual a b***. Para ciertas ecuaciones esto es imposible. Paradójicamente, la mejor manera de entender el caso bueno es estudiando el caso malo. Por consiguiente, consideraremos la geometría, justo cuando falla, en el **caso singular**.

Representación por renglón: intersección de planos ***Representación por columna***: combinación de columnas

El caso singular

Suponga que nuevamente estamos en tres dimensiones, y que tres planos en la representación por renglón *no se cortan*. ¿Qué puede estar mal? Una posibilidad es que dos planos pueden ser paralelos. Las ecuaciones $2u + v + w = 5$ y $4u + 2v + 2w = 11$ son inconsistentes, y planos paralelos no dan solución (en la figura 1.5a se muestra una vista del extremo). En dos dimensiones, la única posibilidad de falla la constituyen las rectas paralelas. Sin embargo, tres planos en tres dimensiones pueden estar en problemas sin ser paralelos.

La dificultad más frecuente se muestra en la figura 1.5b. A partir de la vista del extremo, los planos forman un triángulo. Cada par de planos se corta en una recta, y estas

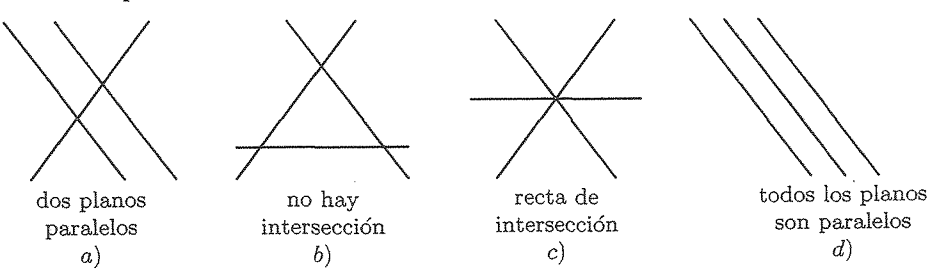


Figura 1.5 Casos singulares: no hay solución para *a*), *b*), o *d*), una infinidad de soluciones para *c*).

rectas son paralelas. El tercer plano no es paralelo a los otros planos, pero es paralelo a su línea de intersección. Esto corresponde a un sistema singular con $b = (2, 5, 6)$:

$$\begin{aligned} u + v + w &= 2 \\ 2u \quad + 3w &= 5 \\ 3u + v + 4w &= 6. \end{aligned} \qquad (3)$$

No hay soluciones, como en la figura 1.5b

Sumados, los dos primeros miembros izquierdos son iguales al tercero. En el miembro derecho falla eso: $2 + 5 \neq 6$. La ecuación 1 menos la ecuación 2 menos la ecuación 3 es la afirmación imposible $0 = 1$. Así, las ecuaciones son ***inconsistentes***, como la eliminación gaussiana descubre sistemáticamente.

Otro sistema singular, próximo a éste, tiene una **infinidad de soluciones**. Una vez que el 6 en la última ecuación se vuelve 7, las tres ecuaciones se combinan para dar $0 = 0$. Así, la tercera ecuación es la suma de las dos primeras. En ese caso, los tres planos tienen toda una *recta en común* (véase la figura 1.5*c*). Al cambiar los miembros derechos, los planos de la figura 1.5*b* se moverán en sentido paralelo a sí mismos, y para $b = (2, 5, 7)$, repentinamente la figura es diferente. El plano inferior se movió para encontrar a los otros, y hay una recta de soluciones. El problema 1.5*c* sigue siendo singular, pero ahora adolece de ***demasiadas soluciones***, en vez de tener unas cuantas.

El caso extremo lo constituyen tres planos paralelos. Para la mayor parte de miembros derechos no hay solución (véase la figura 1.5*d*). Para miembros derechos especiales (¡como $b = (0, 0, 0)$!), hay todo un plano de soluciones, ya que los tres planos paralelos se mueven para convertirse en el mismo.

¿Qué ocurre con la ***representación por columna*** cuando el sistema es singular? Debe estar mal, aunque la pregunta es cómo. En el miembro izquierdo de las ecuaciones sigue habiendo tres columnas, y se intenta combinarlas para obtener b. Se queda con la ecuación (3):

Caso singular: representación por columna
Tres columnas en el mismo plano
Fácil de resolver sólo para b en ese plano

$$u\begin{bmatrix}1\\2\\3\end{bmatrix} + v\begin{bmatrix}1\\0\\1\end{bmatrix} + w\begin{bmatrix}1\\3\\4\end{bmatrix} = b. \qquad (4)$$

Para $b = (2, 5, 7)$ era posible esto; para $b = (2, 5, 6)$ no lo era. La razón de esto es que ***estas tres columnas están en un plano***. Entonces cualquier combinación también está en el plano (que pasa por el origen). Si el vector b no está en ese plano, ninguna solución es posible (véase la figura 1.6). Éste es por mucho el evento más probable; un sistema singular

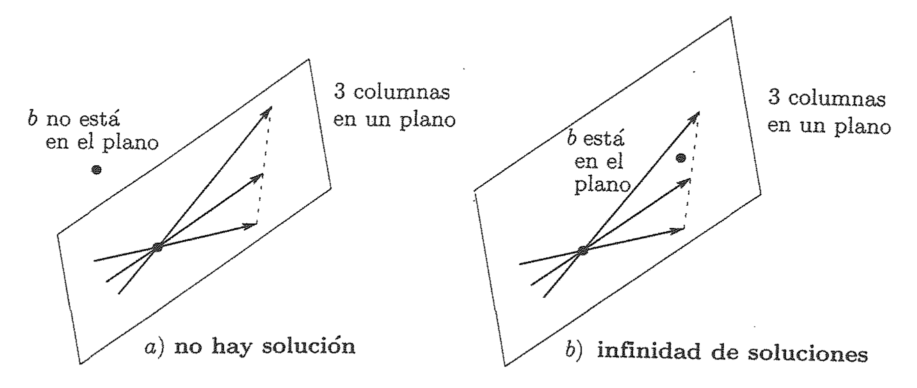


Figura 1.6 Casos singulares: b fuera o dentro del plano con todas las tres columnas.

en general no tiene solución. Sin embargo, hay una posibilidad de que b *esté* en el plano de las columnas. En ese caso hay demasiadas soluciones; las tres columnas pueden combinarse en una *infinidad de formas* para producir b. Esa representación por columna de la figura 1.6*b* corresponde a la representación por renglón de la figura 1.5*c*.

¿Cómo se puede saber que las tres columnas están en el mismo plano? Una respuesta consiste en encontrar una combinación de las columnas cuya suma sea cero. Después de algunos cálculos, esta combinación es $u = 3$, $v = -1$, $w = -2$. Tres veces la columna 1 es igual a dos veces la columna 2 más dos veces la columna 3. La columna 1 está en el plano de las columnas 2 y 3. Sólo dos columnas son independientes.

El vector $b = (2, 5, 7)$ está en ese plano de las columnas: es la columna 1 más la columna 3, de modo que $(1, 0, 1)$ es una solución. *Es posible sumar cualquier múltiplo de la combinación* $(3, -1, -2)$ *que produzca* $b = 0$. Así, hay toda una recta de soluciones, como se sabe a partir de la representación por renglón.

La verdad es que *se sabía* que las columnas deben combinarse para obtener cero, ya que eso ocurría con los renglones. Este hecho pertenece a las matemáticas, no a los cálculos, y sigue siendo verdadero en la dimensión n. ***Si los n planos no tienen ningún punto en común, o comparten una infinidad de puntos, entonces las n columnas están en el mismo plano.***

Si la representación por renglón falla, entonces también falla la representación por columna. Esto hace la diferencia entre el capítulo 1 y el capítulo 2. En el capítulo 1 se estudia el problema más importante: el caso *no singular*, donde hay una solución que es necesario encontrar. En el capítulo 2 se estudia el caso general, donde puede haber muchas soluciones o ninguna. En ninguno de estos dos casos es posible continuar sin tener una notación (*notación matricial*), y un algoritmo (*eliminación*) idóneos. Después de los siguientes ejercicios se abordará la eliminación.

Conjunto de problemas 1.2

1. Para las ecuaciones $x + y = 4$, $2x - 2y = 4$, trace la representación por renglón (dos rectas que se cortan), y la representación por columna (combinación de dos columnas igual al vector columna $(4, 4)$ en el miembro derecho).

2. Resuelva lo siguiente para encontrar una combinación de las columnas que sea igual a b:

Sistema triangular
$$\begin{aligned} u - v - w &= b_1 \\ v + w &= b_2 \\ w &= b_3. \end{aligned}$$

3. (Recomendado) Describa la intersección de los tres planos $u + v + w + z = 6$ y $u + w + z = 4$ y $u + w = 2$ (todos en el espacio tetradimensional). ¿Es una recta, un punto o un conjunto vacío? ¿Cuál es la intersección si se incluye el cuarto plano $u = -1$? Encuentre una cuarta ecuación que deje la situación sin solución.

4. Trace las tres rectas siguientes, y decida si las ecuaciones son de fácil solución:

Sistema de 3 por 2
$$\begin{aligned} x + 2y &= 2 \\ x - y &= 2 \\ y &= 1. \end{aligned}$$

¿Qué ocurre si todos los miembros izquierdos son cero? ¿Hay alguna opción diferente de cero de miembros derechos que permita que las tres rectas se cortan en el mismo punto?

5. Encuentre dos puntos en la recta de intersección de los tres planos $t = 0$ y $z = 0$ y $x + y + z + t = 1$ en el espacio de 4 dimensiones.

6. Cuando $b = (2, 5, 7)$, encuentre una solución (u, v, w) de la ecuación (4) distinta de la solución $(1, 0, 1)$, mencionada en el texto.

7. Proporcione dos miembros derechos más, aparte de $b = (2, 5, 7)$ para los cuales la ecuación (4) pueda resolverse. Proporcione dos miembros derechos más, aparte de $b = (2, 5, 6)$ para los cuales la ecuación (4) no pueda resolverse.

8. Explique por qué el sistema

$$\begin{aligned} u + v + w &= 2 \\ u + 2v + 3w &= 1 \\ v + 2w &= 0 \end{aligned}$$

es singular, encontrando una combinación de las tres ecuaciones que produzca $0 = 1$. ¿Qué valor debe sustituirse en el último cero del miembro derecho para que las ecuaciones tengan soluciones, y cuál es una de las soluciones?

9. La representación por columna del ejercicio anterior (sistema singular) es

$$u \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + v \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + w \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = b.$$

Demuestre que las tres columnas de la izquierda están en el mismo plano, expresando la tercera columna como una combinación de las dos primeras. ¿Cuáles son las soluciones (u, v, w) si b es el vector cero $(0, 0, 0)$?

10. (Recomendado) ¿Bajo qué condición sobre y_1, y_2, y_3 los puntos $(0, y_1), (1, y_2), (2, y_3)$ están en una línea recta?

11. Es cierto que la solución de las siguientes ecuaciones es $x = y = 0$. ¿Para qué valores de a hay toda una recta de soluciones?

$$\begin{aligned} ax + 2y &= 0 \\ 2x + ay &= 0 \end{aligned}$$

12. Empezando con $x + 4y = 7$, encuentre la ecuación de la recta paralela que pasa por $x = 0, y = 0$. Encuentre la ecuación de otra recta que corta a la primera en $x = 3, y = 1$.

Los problemas 13 a 15 son un repaso de las representaciones por renglón y por columna.

13. Trace las dos representaciones en dos planos para las ecuaciones $x - 2y = 0, x + y = 6$.

14. Para dos ecuaciones lineales en tres incógnitas x, y, z, la representación por renglón muestra (2 o 3) (rectas o planos) en un espacio (bi o tri) dimensional. La representación por columna es en un espacio (bi o tri) dimensional. Las soluciones normalmente están en un ______.

15. Para cuatro ecuaciones lineales en dos incógnitas x y y, la representación por renglón muestra cuatro ______. La representación por columna está en un espacio ______ dimensional. Las ecuaciones no tienen solución, a menos que el vector del miembro derecho sea una combinación de ______.

16. Encuentre un punto con $z = 2$ en la recta de intersección de los planos $x + y + 3z = 6$ y $x - y + z = 4$. Encuentre el punto con $z = 0$ y un tercer punto a la mitad entre los dos puntos anteriores.

17. La primera de las siguientes ecuaciones más la segunda es igual a la tercera:

$$\begin{aligned} x + y + z &= 2 \\ x + 2y + z &= 3 \\ 2x + 3y + 2z &= 5. \end{aligned}$$

Los dos primeros planos se encuentran a lo largo de una recta. El tercer plano contiene a esta recta, ya que si x, y, z satisfacen las dos primeras ecuaciones, entonces también ______. Las ecuaciones tienen una infinidad de soluciones (toda la recta **L**). Encuentre las tres soluciones.

18. Mueva el tercer plano en el problema 17 hasta un plano paralelo $2x + 3y + 2z = 9$. Ahora, las tres ecuaciones no tienen solución; *¿por qué*? Los dos primeros planos se encuentran a lo largo de la recta **L**, pero el tercero no ______ esa recta.

19. En el problema 17, las columnas son $(1, 1, 2)$ y $(1, 2, 3)$ y $(1, 1, 2)$. Este es un "caso singular" porque la tercera columna es ______. Encuentre dos combinaciones de las columnas que proporcionen $b = (2, 3, 5)$. Esto sólo es posible para $b = (4, 6, c)$ si $c =$ ______.

20. Normalmente, 4 "planos" en el espacio tetradimensional se cortan en un ______. Normalmente, 4 vectores columna en el espacio de 4 dimensiones pueden combinarse para producir b. ¿Qué combinación de $(1, 0, 0, 0), (1, 1, 0, 0), (1, 1, 1, 0), (1, 1, 1, 1)$ produce $b = (3, 3, 3, 2)$? ¿Cuáles son las 4 ecuaciones que está resolviendo para x, y, z, t?

21. Cuando la ecuación 1 se suma a la ecuación 2, ¿cuál de las siguientes opciones cambia: los planos en la representación por renglón, la representación por columna, la matriz de coeficientes, la solución?

22. Si (a, b) es un múltiplo de (c, d) con $abcd \neq 0$, *demuestre que* (a, c) *es un múltiplo de* (b, d). Esto es sorprendentemente importante: denomínela pregunta de desafío. Primero puede usar números para ver cómo están relacionados a, b, c y d. La pregunta lleva a:

Si $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ tiene renglones dependientes, entonces tiene columnas dependientes.

23. En estas ecuaciones, la tercera columna (que multiplica a w) es *la misma* que el miembro derecho b. La forma en columna de las ecuaciones, ¿qué solución para (u, v, w) proporciona *de inmediato*?

$$\begin{aligned} 6u + 7v + 8w &= 8 \\ 4u + 5v + 9w &= 9 \\ 2u - 2v + 7w &= 7. \end{aligned}$$

1.3 UN EJEMPLO DE ELIMINACIÓN GAUSSIANA

La forma de entender la eliminación es por medio de un ejemplo. Se empieza en tres dimensiones:

$$\textbf{Sistema original} \qquad \begin{aligned} 2u + v + w &= 5 \\ 4u - 6v \qquad &= -2 \\ -2u + 7v + 2w &= 9. \end{aligned} \qquad (1)$$

El problema consiste en encontrar los valores incógnitos de u, v, y w, de modo que se aplicará eliminación gaussiana. (Gauss es reconocido como el más grande de los matemáticos, aunque no ciertamente debido a este invento, que quizá le llevó 10 minutos. Irónicamente, es el concepto más frecuentemente utilizado que lleva su nombre.) El método em-

pieza por ***restar múltiplos de la primera ecuación a las otras ecuaciones***. El objetivo es *eliminar u de las dos últimas ecuaciones*. Para lograr este objetivo se requiere lo siguiente:

a) restar 2 veces la primera ecuación de la segunda
b) restar –1 vez la primera ecuación de la tercera.

Sistema equivalente

$$\begin{aligned} 2u + v + w &= 5 \\ -8v - 2w &= -12 \\ 8v + 3w &= 14. \end{aligned} \qquad (2)$$

El coeficiente 2 es el ***primer pivote***. La eliminación consiste en dividir constantemente el primer pivote entre los números que están abajo de él, con la finalidad de encontrar los multiplicadores idóneos.

El pivote de la **segunda etapa de la eliminación** es –8. Ahora se ignora la primera ecuación. Un múltiplo de la segunda ecuación se restará de las ecuaciones que quedan (en este caso sólo queda la tercera) con la finalidad de eliminar v. La segunda ecuación se suma a la tercera o, en otras palabras,

c) se resta –1 vez la segunda ecuación de la tercera.

Ahora el proceso de eliminación está completo, por lo menos en la dirección "hacia adelante":

Sistema triangular

$$\begin{aligned} 2u + v + w &= 5 \\ -8v - 2w &= -12 \\ 1w &= 2. \end{aligned} \qquad (3)$$

Este sistema se resuelve hacia atrás, de abajo arriba. La última ecuación da $w = 2$. Al sustituir en la segunda ecuación, se encuentra $v = 1$. Luego, la primera ecuación da $u = 1$. Este proceso se denomina ***sustitución hacia atrás***.

Para repetir: con la eliminación hacia adelante se obtuvieron los pivotes $2, -8, 1$. En este método se restan múltiplos de cada renglón de los renglones de abajo para llegar al sistema "triangular" (3), que se resuelve en orden inverso. Luego, cada nuevo valor calculado se sustituye en las ecuaciones restantes.

Observación Una forma aceptable de escribir los pasos de la eliminación hacia adelante es incluir el miembro derecho como una columna adicional. No es necesario copiar u y v y w y $=$ en cada paso, por lo que se trabaja con lo mínimo indispensable:

$$\begin{bmatrix} 2 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{bmatrix} \to \begin{bmatrix} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 8 & 3 & 14 \end{bmatrix} \to \begin{bmatrix} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 0 & 1 & 2 \end{bmatrix}.$$

Al final se llega al sistema triangular, que ya está listo para la sustitución hacia atrás. Quizá el lector prefiera esta disposición, que garantiza que las operaciones en el miembro izquierdo de la ecuación también se realizan en el miembro derecho, ya que *ambos miembros están juntos ahí*.

En un problema más grande, la eliminación hacia adelante requiere más esfuerzo. Se usan múltiplos de la primera ecuación para producir ceros abajo del primer pivote. Luego, la segunda columna se limpia abajo del segundo pivote. El paso hacia adelante se finaliza cuando el sistema es triangular; la ecuación n sólo contiene a la última incógnita multiplicada por el último pivote.

La sustitución hacia atrás produce la solución completa en orden opuesto: se empieza con la última incógnita, luego se resuelve de la siguiente hasta la última, terminando con la primera.

Por definición, ***los pivotes no pueden ser cero***. Es necesario dividir entre ellos.

La falla de la eliminación

¿En qué circunstancias es posible que falle el proceso? Algo *debe* estar mal en el caso singular, y algo *puede* estar mal en el caso no singular. Esto podría parecer algo prematuro; después de todo, apenas se ha logrado tener funcionando al algoritmo. Sin embargo, la posibilidad de falla ilumina al método en sí.

La respuesta es: con un conjunto completo de n pivotes, sólo hay una solución. El sistema es no singular, y se resuelve por eliminación hacia adelante y sustitución hacia atrás. Pero ***si en una posición pivote aparece un cero***, es necesario detener la eliminación, ya sea temporal o permanentemente. El sistema puede o no ser singular.

Si el primer coeficiente es cero, en la esquina superior izquierda, la eliminación de u de las otras ecuaciones es imposible. Lo mismo es cierto en toda etapa intermedia. Observe que en una posición pivote puede aparecer un cero, aun si el coeficiente original en ese sitio no era cero. En términos generales, ***no se sabe si aparecerá un cero sino hasta que se intenta***, al realizar en verdad el proceso de eliminación.

En muchos casos este problema puede restablecerse, por lo que la eliminación puede continuar. Un sistema así sigue siendo no singular; es sólo el algoritmo lo que requiere reparación. En otros casos es inevitable la falla. Estos sistemas incurables son singulares; no tienen solución o tienen una infinidad de éstas, por lo que no es posible encontrar un conjunto completo de pivotes.

Ejemplo 1 No singular (restablecido al intercambiar las ecuaciones 2 y 3)

$$\begin{aligned} u + v + w &= _ \\ 2u + 2v + 5w &= _ \\ 4u + 6v + 8w &= _ \end{aligned} \quad \longrightarrow \quad \begin{aligned} u + v + w &= _ \\ 3w &= _ \\ 2v + 4w &= _ \end{aligned} \quad \longrightarrow \quad \begin{aligned} u + v + w &= _ \\ 2v + 4w &= _ \\ 3w &= _ \end{aligned}$$

El sistema es triangular, y puede resolverse con sustitución hacia atrás.

Ejemplo 2 Singular (caso incurable)

$$\begin{aligned} u + v + w &= _ \\ 2u + 2v + 5w &= _ \\ 4u + 4v + 8w &= _ \end{aligned} \quad \longrightarrow \quad \begin{aligned} u + v + w &= _ \\ 3w &= _ \\ 4w &= _ \end{aligned}$$

No existe ningún intercambio de ecuaciones que pueda evitar el cero en la segunda posición pivote. Las ecuaciones mismas pueden ser o no fáciles de resolver. Si las dos últimas ecuaciones son $3w = 6$ y $4w = 7$, no hay solución. Si ocurre que estas dos ecuaciones son consistentes, como con $3w = 6$ y $4w = 8$, entonces este caso singular tiene una infinidad de soluciones. Se sabe que $w = 2$, pero la primera ecuación no puede decidir ambas u y v.

En la sección 1.5 se abordarán los intercambios de renglón cuando el sistema es no singular. Así, los intercambios producen un conjunto completo de pivotes. En el capítulo 2 se estudia el caso singular. El $3w$ aún puede eliminar a $4w$, por lo que el segundo pivote es 3. (No habrá un tercer pivote). Por ahora se confía en que todos los n elementos pivote son diferentes de cero, sin cambiar el orden de las ecuaciones. Éste es el mejor caso, que será continuado.

El costo de la eliminación

La otra pregunta es muy práctica. *¿Cuántas operaciones aritméticas requiere la eliminación para n ecuaciones en n incógnitas?* Si n es grande, una computadora puede realizar el proceso de eliminación. Debido a que se conocen todos los pasos, debe ser posible poder pronosticar el número de operaciones.

Por el momento se ignorarán los miembros derechos de las ecuaciones, y sólo se contarán las operaciones a la izquierda. Estas operaciones son de dos clases. Se divide entre el pivote para encontrar qué múltiplo (por ejemplo ℓ) de la ecuación pivote debe restarse. Una vez que se realiza esta sustracción, continuamente se encuentra una combinación "multiplicar-restar"; los términos de la ecuación pivote se multiplican por ℓ, y luego se restan de otra ecuación.

Suponga que cada división, y cada multiplicación-sustracción se denomina una operación. En la columna 1, *se requieren n operaciones por cada cero que se obtiene*: una para encontrar el múltiplo ℓ, y la otra para encontrar los nuevos elementos a lo largo del renglón. Abajo del primer renglón hay $n-1$ renglones, de modo que la primera etapa de la eliminación requiere $n(n-1) = n^2 - n$ operaciones. (Otra forma de llegar a $n^2 - n$ es ésta: *es necesario cambiar todos los n^2 elementos, excepto los n en el primer renglón*). Las etapas posteriores son más rápidas porque las ecuaciones son más cortas.

Cuando la eliminación se realiza con k ecuaciones, para limpiar la columna que está abajo del pivote se requieren sólo $k^2 - k$ operaciones, por el mismo razonamiento que se aplicó en la primera etapa, cuando k era igual a n. Al reunir todo lo anterior, se encuentra que el número total de operaciones es la sumatoria de $k^2 - k$ sobre todos los valores k desde 1 hasta n:

Miembro izquierdo $$(1^2 + \cdots + n^2) - (1 + \cdots + n) = \frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)}{2} = \frac{n^3 - n}{3}.$$

Estas fórmulas son normales para encontrar las sumatorias de los n primeros números y los n primeros cuadrados. Al sustituir $n = 1$ y $n = 2$ y $n = 100$ en la fórmula $\frac{1}{3}(n^3 - n)$, la eliminación hacia adelante puede no requerir ningún paso, requerir dos pasos o requerir alrededor de un millón de pasos:

> Si n no es grande en absoluto, ***una buena estimación para el número de operaciones es $\frac{1}{3}n^3$.***

Si el tamaño se duplica, y pocos de los coeficientes son cero, entonces el costo se multiplica por 8.

La sustitución hacia atrás es considerablemente más rápida. La última incógnita se encuentra en sólo una operación (una división entre el último pivote). Para encontrar la antepenúltima incógnita se requieren dos operaciones, y así sucesivamente. Entonces, el total para la sustitución hacia atrás es $1 + 2 + \cdots + n$.

La eliminación hacia adelante también actúa sobre el miembro derecho (restando los mismos múltiplos que en la izquierda con la finalidad de preservar las ecuaciones correctas). Empieza con $n-1$ sustracciones de la primera ecuación. ***Junto con el miembro derecho es responsable de n^2 operaciones***: mucho menos que las $n^3/3$ a la izquierda. El total para la eliminación hacia adelante y la sustitución hacia atrás es

Miembro derecho $$[(n-1) + (n-2) + \cdots + 1] + [1 + 2 + \cdots + n] = n^2.$$

Hace 30 años, casi cualquier matemático hubiera conjeturado que un sistema general de orden n no podía resolverse con mucho menos que $n^3/3$ multiplicaciones. (Incluso había teoremas para demostrarlo, aunque no tomaban en cuenta a todos los métodos posibles). Sorprendentemente, se ha demostrado que esa conjetura es errónea. ¡*En la actualidad existe un método que requiere sólo $Cn^{\log_2 7}$ multiplicaciones*! Depende de un simple hecho: parece que dos combinaciones de dos vectores en el espacio bidimensional requieren 8 multiplicaciones, aunque es posible hacerlas en 7. Lo anterior redujo el exponente de $\log_2 8$, que es 3, a $\log_2 7 \approx 2.8$. Este descubrimiento provocó bastante actividad para encontrar la potencia mínima posible de n. El exponente finalmente cayó (en IBM) debajo de 2.376. Afortunadamente para la eliminación, la constante C es tan grande y el código es tan difícil que el nuevo método es bastante (o completamente) interesante desde el punto de vista teórico. El problema más reciente es el costo con *muchos procesadores en paralelo*.

Conjunto de problemas 1.3

Los problemas 1 a 9 son sobre eliminación en sistemas de 2 por 2.

1. ¿Qué múltiplo de ℓ de la ecuación 1 debe restarse de la ecuación 2?

$$\begin{aligned} 2x + 3y &= 1 \\ 10x + 9y &= 11. \end{aligned}$$

Después de este paso de eliminación, escriba el sistema triangular superior e identifique los dos pivotes. Los números 1 y 11 no afectan tales pivotes.

2. Resuelva por sustitución hacia atrás el sistema triangular del problema 1, y antes que x. Compruebe que x multiplicado por (2, 10) más y por (3, 9) es igual a (1, 11). Si el miembro derecho cambia a (4, 44), ¿cuál es la nueva solución?

3. ¿Qué múltiplo de la ecuación 2 debe *restarse* de la ecuación 3?

$$\begin{aligned} 2x - 4y &= 6 \\ -x + 5y &= 0. \end{aligned}$$

Después de este paso de eliminación, resuelva el sistema triangular. Si el miembro derecho cambia a $(-6, 0)$, ¿cuál es la nueva solución?

4. ¿Qué múltiplo de ℓ de la ecuación 1 debe restarse de la ecuación 2?

$$\begin{aligned} ax + by &= f \\ cx + dy &= g. \end{aligned}$$

El primer pivote es a (se supone que es cero). ¿Qué fórmula para el segundo pivote produce la eliminación? ¿Qué es y? El segundo pivote falta cuando $ad = bc$.

5. Escoja un miembro derecho que no proporcione ninguna solución, y un miembro derecho que no proporcione una infinidad de soluciones. ¿Cuáles son dos de estas soluciones?

$$\begin{aligned} 3x + 2y &= 10 \\ 6x + 4y &= \underline{\quad}. \end{aligned}$$

6. Escoja un coeficiente b que haga singular este sistema. Luego, escoja un miembro derecho g que lo haga fácil de resolver. Encuentre dos soluciones en ese caso singular.

$$\begin{aligned} 2x + by &= 16 \\ 4x + 8y &= g. \end{aligned}$$

7. ¿Para qué números a la eliminación falla *a*) permanentemente, y *b*) temporalmente?

$$\begin{aligned} ax + 3y &= -3 \\ 4x + 6y &= 6. \end{aligned}$$

Resuelva para x y y después de fijar la segunda falla por un intercambio de renglón.

8. ¿Para cuáles tres números k falla la eliminación? ¿Cuál es fijada por un intercambio de renglón? En cada caso, el número de soluciones, ¿es 0 o 1 o ∞?

$$\begin{aligned} kx + 3y &= 6 \\ 3x + ky &= -6. \end{aligned}$$

9. ¿Qué prueba sobre b_1 y b_2 decide si estas dos ecuaciones permiten una solución? ¿Cuántas soluciones tienen? Trace la representación por columna.

$$\begin{aligned} 3x - 2y &= b_1 \\ 6x - 4y &= b_2. \end{aligned}$$

Los problemas 10 a 19 estudian la eliminación en sistemas de 3 por 3 (y la falla posible).

10. Reduzca el siguiente sistema a forma triangular superior, mediante dos operaciones en renglones:

$$\begin{aligned} 2x + 3y + z &= 8 \\ 4x + 7y + 5z &= 20 \\ -2y + 2z &= 0. \end{aligned}$$

Identifique los pivotes. Resuelva por sustitución hacia atrás para z, y, x.

11. Aplique eliminación (identifique los pivotes) y sustitución hacia atrás para resolver

$$\begin{aligned} 2x - 3y \phantom{{}+z} &= 3 \\ 4x - 5y + z &= 7 \\ 2x - y - 3z &= 5. \end{aligned}$$

Enumere las tres operaciones en renglones: Restar _____ veces el renglón _____ del renglón _____.

12. ¿Cuál número d obliga a un intercambio de renglón, y cuál es el sistema triangular (no singular) para ese d? ¿Cuál d hace singular a este sistema (no el tercer pivote)?

$$\begin{aligned} 2x + 5y + z &= 0 \\ 4x + dy + z &= 2 \\ y - z &= 3. \end{aligned}$$

13. ¿Cuál número d lleva posteriormente a un intercambio de renglón? ¿Cuál número b lleva posteriormente a un pivote faltante? En ese caso singular, encuentre una solución x, y, z diferente de cero.

$$\begin{aligned} x + by \phantom{{}-z} &= 0 \\ x - 2y - z &= 0 \\ y + z &= 0. \end{aligned}$$

14. *a*) Escriba un sistema de 3 por 3 que requiera dos intercambios de renglones para alcanzar una forma triangular, y una solución.
b) Escriba un sistema de 3 por 3 que requiera un intercambio de renglones para funcionar, y que falle después.

15. Si los renglones 1 y 2 son los mismos, ¿hasta dónde es posible llegar con la eliminación (permitiendo intercambio de renglones)? Si las columnas 1 y 2 son las mismas, ¿cuál pivote falta?

$$\begin{aligned} 2x - y + z &= 0 \\ 2x - y + z &= 0 \\ 4x + y + z &= 2 \end{aligned} \qquad \begin{aligned} 2x + 2y + z &= 0 \\ 4x + 4y + z &= 0 \\ 6x + 6y + z &= 2. \end{aligned}$$

16. Escriba un ejemplo de 3 por 3 que tenga 9 coeficientes distintos en el miembro izquierdo, pero que los renglones 2 y 3 se vuelvan cero en la eliminación. ¿Cuántas soluciones tiene este sistema con $b = (1, 10, 100)$, y cuántas tiene con $b = (0, 0, 0)$?

17. ¿Cuál número q hace singular al siguiente sistema, y con qué miembro derecho t el sistema tiene una infinidad de soluciones? Encuentre la solución que tiene $z = 1$.

$$\begin{aligned} x + 4y - 2z &= 1 \\ x + 7y - 6z &= 6 \\ 3y + qz &= t. \end{aligned}$$

18. (Recomendado) Es imposible que un sistema de ecuaciones lineales tenga exactamente dos soluciones. *Explique por qué.*
a) Si (x, y, z) y (X, Y, Z) son dos soluciones, ¿cuál es otra?
b) Si 25 planos se encuentran en dos puntos, ¿dónde más se encuentran?

19. Puede ocurrir que tres planos fallen en tener un punto de intersección, cuando ningún par de ellos es paralelo. El sistema es singular si el renglón 3 de A es una _____ de los primeros dos renglones. Encuentre una tercera ecuación que no sea posible resolver si $x + y + z = 0$ y $x - 2y - z = 1$.

En los problemas 20 a 22 se abordan sistemas de 4 por 4, y de n por n.

20. Encuentre los pivotes, y las soluciones de las cuatro siguientes ecuaciones:

$$\begin{aligned} 2x + y \phantom{{}+z+t} &= 0 \\ x + 2y + z \phantom{{}+t} &= 0 \\ y + 2z + t &= 0 \\ z + 2t &= 5. \end{aligned}$$

21. Si el problema 20 se extiende siguiendo el patrón 1, 2, 1 o el patrón −1, 2, −1, ¿cuál es el quinto pivote? ¿Cuál es el n-ésimo pivote?

22. Aplique eliminación y sustitución hacia atrás para resolver

$$\begin{aligned} 2u + 3v \phantom{{}+w} &= 0 \\ 4u + 5v + w &= 3 \\ 2u - v - 3w &= 5. \end{aligned}$$

¿Cuáles son los pivotes? Enumere las tres operaciones en las que el múltiplo de un renglón se resta de otro renglón.

23. Para el sistema

$$\begin{aligned} u + v + w &= 2 \\ u + 3v + 3w &= 0 \\ u + 3v + 5w &= 2, \end{aligned}$$

¿Cuál es el sistema triangular después de la eliminación hacia adelante, y cuál es la solución?

24. Resuelva el siguiente sistema, y encuentre los pivotes cuando

$$\begin{aligned} 2u - v \quad\quad &= 0 \\ -u + 2v - w \quad &= 0 \\ -v + 2w - z &= 0 \\ -w + 2z &= 5. \end{aligned}$$

Es posible llevar el miembro derecho como una quinta columna (y así omitir escribir u, v, w, z hasta la solución al final).

25. Aplique eliminación al sistema

$$\begin{aligned} u + v + w &= -2 \\ 3u + 3v - w &= 6 \\ u - v + w &= -1. \end{aligned}$$

Cuando en la posición pivote aparezca un cero, intercambie esa ecuación por la que está abajo, y continúe. *¿Qué coeficiente de v en la tercera ecuación, en lugar del -1 actual, haría imposible continuar, y forzar a una falla en la eliminación?*

26. Resuelva por eliminación el siguiente sistema de ecuaciones

$$\begin{aligned} x - y &= 0 \\ 3x + 6y &= 18. \end{aligned}$$

Trace una gráfica que representa cada ecuación como una línea recta en el plano $x - y$; las rectas se cortan en la solución. También, agregue una recta más: la gráfica de una nueva segunda ecuación que aparece después de la eliminación.

27. Encuentre tres valores de a para los cuales falle la eliminación, temporal o permanentemente, en

$$\begin{aligned} au + v &= 1 \\ 4u + av &= 2. \end{aligned}$$

La falla en el primer paso puede establecerse intercambiando renglones, pero en el último paso no hay falla.

28. Falso o verdadero:

a) Si la tercera ecuación inicia con un coeficiente cero (empieza con $0u$), entonces ningún múltiplo de la ecuación 1 se resta de la ecuación 3.

b) Si el segundo coeficiente de la tercera ecuación es cero (contiene a $0v$), entonces ningún múltiplo de la ecuación 2 se resta de la ecuación 3.

c) Si la tercera ecuación contiene a $0u$ y a $0v$, entonces ningún múltiplo de la ecuación 1 o de la ecuación 2 se resta de la ecuación 3.

29. (Bastante opcional). Normalmente la multiplicación de dos números complejos

$$(a + ib)(c + id) = (ac - bd) + i(bc + ad)$$

implica las cuatro multiplicaciones por separado ac, bd, bc, ad. Ignorando a i, ¿puede calcular $ac - bd$ y $bc + ad$ con sólo tres multiplicaciones? (Puede realizar sumas, como formar $a + b$ antes de multiplicar, sin ninguna penalización.)

30. Aplique eliminación para resolver

$$\begin{aligned} u + v + w &= 6 \\ u + 2v + 2w &= 11 \\ 2u + 3v - 4w &= 3 \end{aligned} \quad \text{y} \quad \begin{aligned} u + v + w &= 7 \\ u + 2v + 2w &= 10 \\ 2u + 3v - 4w &= 3. \end{aligned}$$

31. ¿Para cuáles tres números a la eliminación fracasa en proporcionar tres pivotes?

$$\begin{aligned} ax + 2y + 3z &= b_1 \\ ax + ay + 4z &= b_2 \\ ax + ay + az &= b_3. \end{aligned}$$

32. Encuentre experimentalmente el tamaño medio (valor absoluto) de los pivotes primero, segundo y tercero para MATLAB con lu(rand(3, 3)). El promedio del primer pivote a partir de abs(A(1, 1)) debe ser 0.5.

1.4 NOTACIÓN MATRICIAL Y MULTIPLICACIÓN DE MATRICES

Con el ejemplo de 3 por 3 es posible escribir por completo todas las ecuaciones. Es posible enumerar los pasos de la eliminación, los cuales restan un múltiplo de una ecuación a otra para llegar a una matriz triangular. Para un sistema grande, esta forma de mantener el rastro de la eliminación sería inútil; se requiere un registro mucho más conciso.

A continuación se presenta la **notación matricial** para describir el sistema original, y la **multiplicación de matrices** para describir las operaciones que lo hacen más sencillo. Observe que en el ejemplo aparecen tres tipos distintos de cantidades:

Nueve coeficientes
Tres incógnitas
Tres miembros derechos

$$\begin{aligned} 2u + v + w &= 5 \\ 4u - 6v \quad\quad &= -2 \\ -2u + 7v + 2w &= 9 \end{aligned} \qquad (1)$$

En el miembro derecho está el vector columna b. En el miembro izquierdo están las incógnitas u, v, w. En el miembro izquierdo también están nueve coeficientes (de los cuales uno es cero). Resulta natural representar las tres incógnitas por medio de un vector:

$$\text{La incógnita es } x = \begin{bmatrix} u \\ v \\ w \end{bmatrix} \qquad \text{La solución es } x = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}.$$

Los nueve coeficientes están en tres renglones y tres columnas, con lo que se obtiene una ***matriz de 3 por 3:***

$$\textbf{Matriz de coeficientes } A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}.$$

A es una matriz *cuadrada*, porque el número de ecuaciones es igual al número de incógnitas. Si hay n ecuaciones en n incógnitas, se tiene una matriz cuadrada de n por n. De manera más general, podría haber m ecuaciones y n incógnitas. En este caso, A es *rectangular*, con m renglones y n columnas. Se dice que es una "matriz de m por n".

Las matrices se suman entre sí, o multiplican por constantes numéricas, exactamente como ocurre con los vectores: elemento por elemento. De hecho, los vectores pueden considerarse como casos especiales de las matrices; *son matrices con una sola columna*. Así como con los vectores, la suma de dos matrices sólo es posible si tienen la misma forma:

Adición $A + B$
Multiplicación $2A$

$$\begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ -3 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 0 & 1 \\ 1 & 6 \end{bmatrix} \qquad 2\begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 6 & 0 \\ 0 & 8 \end{bmatrix}.$$

Multiplicación de una matriz por un vector

Es necesario volver a escribir las tres ecuaciones con las tres incógnitas u, v, w en la forma matricial simplificada $Ax = b$. Escrita completamente, la multiplicación de una matriz por un vector es igual al vector:

Forma matricial $Ax = b$
$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}. \tag{2}$$

El miembro derecho b es el vector columna de los "términos no homogéneos". ***El miembro izquierdo es A por x***. Esta multiplicación se define *exactamente de modo que reproduzca el sistema original*. La primera componente de Ax se obtiene al "multiplicar" el primer renglón de A en el vector columna x:

Renglón multiplicado por columna
$$\begin{bmatrix} 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 2u + v + w \end{bmatrix} = \begin{bmatrix} 5 \end{bmatrix}. \tag{3}$$

La segunda componente del producto Ax es $4u - 6v + 0w$, del segundo renglón de A. La ecuación matricial $Ax = b$ es equivalente a las tres ecuaciones simultáneas en la ecuación (1).

El ***renglón por columna*** es fundamental para todas las multiplicaciones de matrices. A partir de dos vectores produce un solo número. Este número se denomina ***producto interno*** de los dos vectores. En otras palabras, el producto de una matriz de 1 por n (*un vector renglón*) y una matriz de n por 1 (*un vector columna*) es una matriz de 1 por 1:

Producto interno
$$\begin{bmatrix} 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \cdot 1 + 1 \cdot 1 + 1 \cdot 2 \end{bmatrix} = \begin{bmatrix} 5 \end{bmatrix}.$$

Esto confirma que la solución propuesta $x = (1, 1, 2)$ satisface la primera ecuación.

Hay dos formas de multiplicar una matriz A y un vector x. Una forma es multiplicar *renglón por renglón*. Cada renglón de A se combina con x para obtener una componente de Ax. Cuando A tiene tres renglones, se tienen tres productos.

Ax multiplicado por renglones
$$\begin{bmatrix} 1 & 1 & 6 \\ 3 & 0 & 1 \\ 1 & 1 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \cdot 2 + 1 \cdot 5 + 6 \cdot 0 \\ 3 \cdot 2 + 0 \cdot 5 + 3 \cdot 0 \\ 1 \cdot 2 + 1 \cdot 5 + 4 \cdot 0 \end{bmatrix} = \begin{bmatrix} 7 \\ 6 \\ 7 \end{bmatrix}. \tag{4}$$

Así es como suele explicarse Ax, aunque la segunda forma es igualmente importante. De hecho, ¡es más importante! Realiza la multiplicación de *una columna a la vez*. El producto Ax se encuentra de una vez, como ***una combinación de las tres columnas de A:***

Ax multiplicado por columnas
$$2\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + 5\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 0\begin{bmatrix} 6 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ 6 \\ 7 \end{bmatrix}. \tag{5}$$

La respuesta es dos veces la columna 1 más 5 veces la columna 2. Esto corresponde a la "representación por columna" de las ecuaciones lineales. Si el miembro derecho b tiene las componentes 7, 6, 7, entonces la solución tiene componentes 2, 5, 0. Por supuesto, la representación por renglón coincide con aquélla (y finalmente se tienen las mismas multiplicaciones).

La regla de las columnas se utilizará una y otra vez, de modo que para recalcarla se repite a continuación:

> **1A** Todo producto Ax puede encontrarse usando todas las columnas como en la ecuación (5). Por consiguiente, Ax es ***una combinación de las columnas de A***. Los coeficientes son las componentes de x.

Para multiplicar A por x en n dimensiones, se requiere una notación para los elementos individuales en A. *El elemento en el renglón i y en la columna j siempre se denota por* a_{ij}. El primer subíndice proporciona el número de renglón, y el segundo subíndice indica la columna. (En la ecuación (4), a_{21} es 3 y a_{13} es 6.) Si A es una matriz de m por n, entonces el índice i va desde 1 hasta m, ya que hay m renglones, y el índice j va desde 1 hasta n. En total, la matriz tiene mn elementos, y a_{mn} está en la esquina inferior derecha.

Para denotar un vector basta un índice. La j-ésima componente de x se denota x_j. (La multiplicación anterior tenía $x_1 = 2$, $x_2 = 5$, $x_3 = 0$.) Normalmente, x se escribe como un vector columna; es decir, como una matriz de n por 1. Pero algunas veces se escribe en una línea, como en $x = (2, 5, 0)$. Los paréntesis y las comas recalcan que no es una matriz de 1 por 3. Se trata de un vector columna, que sólo se ha escrito temporalmente en forma horizontal.

Para describir el producto Ax se utiliza el símbolo "*sigma*" Σ para sumatorias:

Notación sigma La i-ésima componente de Ax es $\displaystyle\sum_{j=1}^{n} a_{ij}x_j$.

Esta suma se lleva el i-ésimo renglón de A. El índice de la columna j toma cada valor desde 1 hasta n y los resultados se suman: la suma es $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$.

De nuevo se observa que la longitud de los renglones (el número de columnas en A) debe corresponder a la longitud de x. **Una matriz de m por n multiplica un vector n-dimensional** (y produce un vector m-dimensional). Las sumatorias son más sencillas que escribir todo completamente, pero la notación matricial es mejor. (Einstein utilizó "notación tensorial", en la que un índice repetido significa automáticamente una sumatoria. Escribió $a_{ij}x_j$, e incluso $a_i^j x_j$, sin el signo Σ. Como no somos Einstein, mantenemos la Σ.)

Forma matricial de un paso de eliminación

Hasta el momento, se ha utilizado una abreviatura conveniente $Ax = b$ para denotar el sistema de ecuaciones original. Pero, ¿qué hay respecto de las operaciones que se realizan durante la eliminación? En nuestro ejemplo, en el primer paso se restó 2 veces la primera

ecuación de la segunda. En el miembro derecho, 2 veces la primera componente de b se restó de la segunda componente. *Se obtiene el mismo resultado si b se multiplica por esta matriz elemental* (o *matriz de eliminación*):

$$\textbf{Matriz elemental} \qquad E = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Lo anterior se comprueba simplemente obedeciendo la regla para multiplicar una matriz y un vector:

$$Eb = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix} = \begin{bmatrix} 5 \\ -12 \\ 9 \end{bmatrix}.$$

Las componentes 5 y 9 siguen siendo las mismas (debido a los 1, 0, 0 y 0, 0, 1 en los renglones de E). La nueva segunda componente, -12, apareció después del primer paso de eliminación.

Es fácil describir las matrices como E, que llevan los pasos de eliminación por separado. También se observa la "matriz identidad", que no hace absolutamente nada.

1B La *matriz identidad I*, con 1s en su diagonal y 0s en el resto, deja sin cambio a todos los vectores. La *matriz elemental* E_{ij} resta ℓ veces el renglón j del renglón i. Esta E_{ij} contiene a $-\ell$ en el renglón i, columna j.

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ tiene } Ib = b \qquad E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\ell & 0 & 1 \end{bmatrix} \text{ tiene } E_{31}b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 - \ell b_1 \end{bmatrix}.$$

$Ib = b$ es la matriz análoga a multiplicar por 1. Un paso de eliminación típico multiplica por E_{31}. La pregunta importante es: ¿Qué ocurre en el miembro izquierdo de A?

Para preservar la igualdad, es necesario realizar las mismas operaciones en ambos miembros de $Ax = b$. En otras palabras, también el vector Ax debe multiplicarse por la matriz E. Nuestra matriz original E resta 2 veces la primera componente de la segunda. Después de este paso, el nuevo sistema más sencillo (equivalente al anterior) es justamente $E(Ax) = Eb$. Es más sencillo debido al cero que se creó abajo del primer pivote. Es equivalente porque es posible recuperar el sistema original (sumando 2 veces la primera ecuación de regreso a la segunda). Así, ambos sistemas tienen exactamente la misma solución x.

Multiplicación de matrices

Ahora hemos llegado a la pregunta más importante: *¿Cómo se multiplican las matrices?* Hay una pista parcial de la eliminación gaussiana: Se conoce la matriz de coeficientes original A, se conoce la matriz de eliminación E, y se conoce el resultado EA después del paso de eliminación. Deseamos y esperamos que

$$E = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{matrix}\text{multi-}\\ \text{plicado}\\ \text{por}\end{matrix} A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \begin{matrix}\text{propor-}\\ \text{ciona}\end{matrix} EA = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ -2 & 7 & 2 \end{bmatrix}.$$

Dos veces el primer renglón de A se ha restado del segundo renglón. La multiplicación de matrices es consistente con las operaciones por renglones de eliminación. El resultado puede escribirse como $E(Ax) = Eb$, aplicando E a ambos miembros de la ecuación, o como $(EA)x = Eb$. La matriz EA se construye exactamente de modo que estas ecuaciones coincidan, por lo que no se requieren paréntesis:

Multiplicación de matrices (EA multiplicada por x) es igual a (E multiplicada por Ax). Simplemente se escribe EAx.

Esta es toda la cuestión de una "ley asociativa" como $2 \times (3 \times 4) = (2 \times 3) \times 4$. La ley parece tan evidente que es difícil imaginar que podría ser falsa. Pero lo mismo puede decirse de la "ley conmutativa" $2 \times 3 = 3 \times 2$; y para matrices, se tiene que EA no es AE.

Para la multiplicación de matrices hay otro requerimiento. Se sabe cómo multiplicar Ax, una matriz y un vector. La nueva definición debe ser consistente con esta definición. Cuando una matriz B sólo contiene una simple columna x, el producto matriz-matriz AB debe ser idéntico al producto matriz-vector Ax. *Más aún*: Cuando B contiene varias columnas b_1, b_2, b_3, ¡las columnas de AB deben ser Ab_1, Ab_2, Ab_3!

$$\textbf{Multiplicación por columnas} \qquad AB = A\begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & Ab_3 \end{bmatrix}.$$

El primer requerimiento tenía que ver con renglones, y éste concierne a columnas. Un tercer método es describir cada elemento individual en AB y esperar lo mejor. De hecho, sólo hay una regla posible, aunque no estoy seguro de quién la descubrió. Hace funcionar todo. No permite la multiplicación de cualquier par de matrices. Si son cuadradas, deben tener el mismo tamaño. Si son rectangulares, *no* deben tener la misma forma; ***el número de columnas en A debe ser igual al número de renglones en B***. Así, A puede multiplicarse por cada columna de B.

Si A es m por n, y B es n por p, entonces la multiplicación es posible. *El producto AB es una matriz de m por p.* A continuación se encontrará el elemento en el renglón i y columna j de AB.

1C El elemento i, j de AB es el producto interno del i-ésimo renglón de A y la j-ésima columna de B. En la figura 1.7, el elemento 3, 2 de AB proviene del renglón 3 y la columna 2:

$$(AB)_{32} = a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} + a_{34}b_{42}. \tag{6}$$

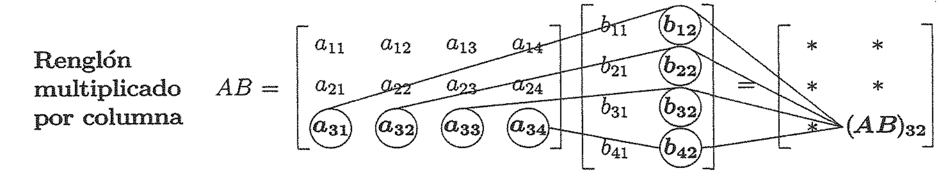


Figura 1.7 Una matriz A de 3 por 4, multiplicada por una matriz B de 4 por 2, es una matriz AB de 3 por 2.

Nota Se escribe AB cuando las matrices no tienen nada especial que hacer con la eliminación. El primer ejemplo era EA debido a la matriz elemental E. Después se tiene PA, o LU, e incluso LDU. La regla para la multiplicación de matrices permanece igual.

Ejemplo 1

$$AB = \begin{bmatrix} 2 & 3 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 5 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 17 & 1 & 0 \\ 4 & 8 & 0 \end{bmatrix}.$$

El elemento 17 es $(2)(1) + (3)(5)$, el producto interno del primer renglón de A y la primera columna de B. El elemento 8 es $(4)(2) + (0)(-1)$, el producto interno del segundo renglón y la segunda columna.

La tercera columna es cero en B, por lo que es cero en AB. B consta de tres columnas una junto a la otra, y A multiplica cada columna por separado. ***Cada columna de AB es una combinación de las columnas de A.*** Así como en la multiplicación matriz-vector, las columnas de A se multiplican por los elementos de B.

Ejemplo 2 **Matriz de intercambio de renglones** $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 7 & 8 \\ 2 & 3 \end{bmatrix}.$

Ejemplo 3 Los 1s en la matriz identidad I dejan igual a la matriz:

Matriz identidad $IA = A$ y $BI = B$.

Importante: La multiplicación AB también puede efectuarse *renglón por renglón*. En el ejemplo 1, el primer renglón de AB utiliza los números 2 y 3 del primer renglón de A. Con estos números se obtiene 2 [renglón 1] + 3 [renglón 2] = [17 1 0]. Exactamente como en la eliminación, donde empezó todo esto, cada renglón de AB ***es una combinación de los renglones de B.***

A continuación se resumen estas tres formas de considerar la multiplicación de matrices.

1D i) Cada elemento de AB es el producto de un ***renglón*** y una ***columna***:

$(AB)_{ij}$ = (renglón i de A) multiplicado por la (columna j de B)

ii) Cada columna de AB es el producto de una ***matriz*** y una **columna**:

columna j de AB = A multiplicada por la (columna j de B)

iii) Cada renglón de AB es el producto de un ***renglón*** y una ***matriz***:

renglón i de AB = (renglón i de A) multiplicado por B

Esto nos recuerda a una propiedad clave de la multiplicación de matrices. Suponga que las formas de tres matrices A, B, C (quizá rectangulares) permiten su multiplicación. Los renglones en A y B se multiplican por las columnas de B y C. Así, la propiedad clave es la siguiente:

1E La multiplicación de matrices es asociativa: $(AB)C = A(BC)$. Se escribe simplemente ABC.

AB multiplicado por C es igual a A multiplicado por BC. Si ocurre que C es simplemente un vector (una matriz con una sola columna), este es el requerimiento $(EA)x = E(Ax)$ ya mencionado. Esto constituye toda la base para las leyes de la multiplicación de matrices. Y si C consta de varias columnas, basta pensar que éstas se encuentran una junto a la otra, y aplicar la misma regla varias veces. Cuando se multiplican varias matrices no se requieren paréntesis.

Queda por mencionar otras dos propiedades: una que posee la multiplicación de matrices, y otra que *no posee*. La primera es:

1F Las operaciones con matrices son distributivas:

$A(B + C) = AB + AC$ y $(B + C)D = BD + CD$.

Por supuesto, las formas de estas matrices deben corresponder correctamente: B y C tienen la misma forma, de modo que es posible sumarlas, y A y D son del tamaño idóneo para la multiplicación por la izquierda y por la derecha. La demostración de esta ley es demasiado tediosa como para escribirla.

La propiedad que no se cumple para la multiplicación de matrices es un poco más interesante:

1G La multiplicación de matrices no es conmutativa. Casi siempre $FE \neq EF$.

Ejemplo 4 Suponga que E resta dos veces la primera ecuación de la segunda. Suponga que F es la matriz del siguiente paso, *sumar el renglón* 1 *al renglón* 3:

$$E = \begin{bmatrix} 1 & 0 & 0 \\ \mathbf{-2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{y} \quad F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mathbf{1} & 0 & 1 \end{bmatrix}.$$

Estas dos matrices conmutan, y el producto efectúa ambos pasos al mismo tiempo:

$$EF = \begin{bmatrix} 1 & 0 & 0 \\ \mathbf{-2} & 1 & 0 \\ \mathbf{1} & 0 & 1 \end{bmatrix} = FE.$$

En cualquier orden, EF o FE, así se cambian los renglones 2 y 3 usando el renglón 1.

Ejemplo 5 Suponga que E es la misma pero que G *suma el renglón* 2 *al renglón* 3. Ahora el orden establece una diferencia. Cuando se aplica E y luego G, el segundo renglón está modificado *antes* de afectar al tercero. Si E se aplica *después* de G, entonces la tercera ecuación no sufre ningún efecto del primero. En el elemento (3, 1) de EG se verá un cero, mientras en GE hay un –2:

$$GE = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \mathbf{1} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ \mathbf{-2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ \mathbf{-2} & 1 & 1 \end{bmatrix} \quad \text{pero} \quad EG = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ \mathbf{0} & 1 & 1 \end{bmatrix}.$$

Así, $EG \neq GE$. Un ejemplo aleatorio mostraría el mismo hecho: la mayor parte de las matrices no conmutan. Aquí las matrices tienen sentido. Hay una razón por la cual $EF = FE$, y otra por la cual $EG \neq GE$. Merece la pena realizar otro paso, para ver lo que ocurre con *las tres matrices de eliminación a la vez*:

$$GFE = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ \mathbf{-1} & 1 & 1 \end{bmatrix} \quad \text{y} \quad EFG = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ \mathbf{1} & 1 & 1 \end{bmatrix}.$$

El producto GFE es el verdadero orden de eliminación. *Es la matriz que lleva la matriz original A a la triangular superior U*. Este hecho se verá de nuevo en la siguiente sección.

La otra matriz, EFG, es más agradable. En ese orden, los números –2 de E y 1 de F y G no fueron alterados. Se escribieron igual en el producto. Es el orden incorrecto para la

eliminación. Pero por fortuna, *se trata del orden correcto para invertir los pasos de eliminación*, lo cual también se verá en la siguiente sección.

Observe que el producto de matrices triangulares inferiores también es triangular inferior.

Conjunto de problemas 1.4

1. Calcule los productos

$$\begin{bmatrix}4&0&1\\0&1&0\\4&0&1\end{bmatrix}\begin{bmatrix}3\\4\\5\end{bmatrix} \quad \text{y} \quad \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}5\\-2\\3\end{bmatrix} \quad \text{y} \quad \begin{bmatrix}2&0\\1&3\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}.$$

Para el tercero, trace los vectores columna (2, 1) y (0, 3). La multiplicación por (1, 1) simplemente suma los vectores (hágalo gráficamente).

2. Trabaje con una columna a la vez para calcular los productos

$$\begin{bmatrix}4&1\\5&1\\6&1\end{bmatrix}\begin{bmatrix}1\\3\end{bmatrix} \quad \text{y} \quad \begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}\begin{bmatrix}0\\1\\0\end{bmatrix} \quad \text{y} \quad \begin{bmatrix}4&3\\6&6\\8&9\end{bmatrix}\begin{bmatrix}\frac{1}{2}\\\frac{1}{3}\end{bmatrix}.$$

3. Encuentre dos productos internos y un producto de matrices:

$$\begin{bmatrix}1&-2&7\end{bmatrix}\begin{bmatrix}1\\-2\\7\end{bmatrix} \quad \text{y} \quad \begin{bmatrix}1&-2&7\end{bmatrix}\begin{bmatrix}3\\5\\1\end{bmatrix} \quad \text{y} \quad \begin{bmatrix}1\\-2\\7\end{bmatrix}\begin{bmatrix}3&5&1\end{bmatrix}.$$

Con el primero se obtiene la longitud del vector (al cuadrado).

4. Si una matriz A de m por n multiplica a un vector n dimensional x, ¿cuántas multiplicaciones por separado están implicadas? ¿Qué ocurre si A multiplica a una matriz B de n por p?

5. Multiplique Ax para encontrar un vector solución x del sistema $Ax =$ *vector cero*. ¿Puede encontrar más soluciones para $Ax = 0$?

$$Ax = \begin{bmatrix}3&-6&0\\0&2&-2\\1&-1&-1\end{bmatrix}\begin{bmatrix}2\\1\\1\end{bmatrix}.$$

6. Escriba las matrices A y B de 2 por 2 cuyos elementos son $a_{ij} = i + j$ y $b_{ij} = (-1)^{i+j}$. Multiplíquelas para encontrar AB y BA.

7. Proporcione ejemplos de 3 por 3 (no sólo la matriz cero) de

a) una matriz diagonal: $a_{ij} = 0$ si $i \neq j$.
b) una matriz simétrica: $a_{ij} = a_{ji}$ para toda i y j.
c) una matriz triangular superior: $a_{ij} = 0$ si $i > j$.
d) una matriz simétrica sesgada: $a_{ij} = -a_{ji}$ para toda i y j.

8. Las siguientes subrutinas, ¿multiplican Ax *por renglones o por columnas*? Empiece con $B(I) = 0$:

```
   DO 10 I = 1,N                      DO 10 J = 1,N
   DO 10 J = 1,N                      DO 10 I = 1,N
10 B(I) = B(I) + A(I,J) * X(J)     10 B(I) = B(I) + A(I,J) * X(J)
```


Las salidas $Bx = Ax$ son las mismas. El segundo código es ligeramente más eficaz en FORTRAN, y mucho más eficaz en una máquina vectorial (el primero cambia los elementos simples $B(I)$, mientras el segundo es capaz de actualizar vectores completos).

9. Si los elementos de A son a_{ij}, use notación de subíndices para escribir

a) El primer pivote.
b) El multiplicador ℓ_{i1} del renglón 1 que debe restarse del renglón i.
c) El nuevo elemento que sustituye a a_{ij} después de la sustracción.
d) El segundo pivote.

10. ¿Falso o verdadero? Cuando sea falso, proporcione un contraejemplo.

a) Si las columnas 1 y 3 de B son las mismas, también las columnas 1 y 3 de AB son las mismas.
b) Si los renglones 1 y 3 de B son los mismos, también los renglones 1 y 3 de AB son los mismos.
c) Si los renglones 1 y 3 de A son los mismos, también los renglones 1 y 3 de AB son los mismos.
d) $(AB)^2 = A^2B^2$.

11. El primer renglón de AB es una combinación lineal de todos los renglones de B. ¿Cuáles son los coeficientes en esta combinación? y ¿cuál es el primer renglón de AB, si

$$A = \begin{bmatrix}2&1&4\\0&-1&1\end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix}1&1\\0&1\\1&0\end{bmatrix}?$$

12. El producto de dos matrices triangulares inferiores es nuevamente triangular inferior (todos sus elementos arriba de la diagonal principal son cero). Confirme esto con un ejemplo de 3 por 3, y luego explique por qué este hecho se deduce a partir de las leyes de la multiplicación de matrices.

13. Por ensayo y error, encuentre ejemplos de matrices de 2 por 2 tales que

a) $A^2 = -I$, donde A sólo tiene elementos reales.
b) $B^2 = 0$, aunque $B \neq 0$.
c) $CD = -DC$, no permitiendo el caso $CD = 0$.
d) $EF = 0$, aunque ningún elemento de E o F es cero.

14. Describa los renglones de EA y las *columnas* de AE si

$$E = \begin{bmatrix}1&7\\0&1\end{bmatrix}.$$

15. Suponga que A conmuta con toda matriz de 2 por 2 ($AB = BA$), y en particular

$$A = \begin{bmatrix}a&b\\c&d\end{bmatrix} \quad \text{conmuta con} \quad B_1 = \begin{bmatrix}1&0\\0&0\end{bmatrix} \quad \text{y} \quad B_2 = \begin{bmatrix}0&1\\0&0\end{bmatrix}.$$

Demuestre que $a = d$ y $b = c = 0$. Si $AB = BA$ para todas las matrices B, entonces A es un múltiplo de la identidad.

16. Sea x el vector columna $(1, 0, \ldots, 0)$. Demuestre que la regla $(AB)x = A(Bx)$ obliga a que la primera columna de AB sea igual a A multiplicada por la primera columna de B.

17. ¿Cuál(es) de la(s) siguiente(s) matrices garantizan ser iguales a $(A + B)^2$?

$$A^2 + 2AB + B^2, \quad A(A + B) + B(A + B), \quad (A + B)(B + A), \quad A^2 + AB + BA + B^2.$$

18. Si A y B son matrices de n por n con todos los elementos iguales a 1, encuentre $(AB)_{ij}$. La notación de la sumatoria transforma el producto AB, y la ley $(AB)C = A(BC)$, en

$$(AB)_{ij} = \sum_k a_{ik}b_{kj} \qquad \sum_j\left(\sum_k a_{ik}b_{kj}\right)c_{jl} = \sum_k a_{ik}\left(\sum_j b_{kj}c_{jl}\right).$$

Calcule ambos miembros de C si también es de n por n, con todo $c_{jl} = 2$.

19. *Una cuarta forma de multiplicar matrices es* ***multiplicando las columnas de A por los renglones de B*****:**

AB = (columna 1)(renglón 1) + $\cdots$ + (columna n)(renglón n) = suma de matrices simples.

Proporcione un ejemplo de 2 por 2 de esta importante regla para la multiplicación de matrices.

20. La matriz que rota el plano x-y en un ángulo θ es

$$A(\theta) = \begin{bmatrix} \cos\theta & -\text{sen}\,\theta \\ \text{sen}\,\theta & \cos\theta \end{bmatrix}.$$

Compruebe que $A(\theta_1)A(\theta_2) = A(\theta_1 + \theta_2)$ a partir de las identidades para $\cos(\theta_1 + \theta_2)$ y $\text{sen}(\theta_1 + \theta_2)$. ¿A qué es igual $A(\theta)$ multiplicada por $A(-\theta)$?

21. Encuentre las potencias A^2, A^3 (A^2 multiplicada por A), y B^2, B^3, C^2, C^3. ¿Qué son A^k, B^k y C^k?

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{y} \quad C = AB = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

Los problemas 22 a 31 son sobre matrices de eliminación.

22. Escriba las matrices de 3 por 3 que producen los siguientes pasos de eliminación:

a) E_{21} resta 5 veces el renglón 1 del renglón 2.

b) E_{32} resta -7 veces el renglón 2 del renglón 3.

c) P intercambia los renglones 1 y 2, y luego los renglones 2 y 3.

23. En el problema 22, al aplicar E_{21} y luego E_{32} a la columna $b = (1, 0, 0)$ se obtiene $E_{32}E_{21}b$ = _____. Al aplicar E_{32} antes que E_{21} se obtiene $E_{21}E_{32}b$ = _____. Cuando se aplica primero E_{32}, el renglón _____ no se ve afectado por el renglón _____.

24. ¿Cuáles son las tres matrices E_{21}, E_{31}, E_{32} que hacen de A una forma triangular U?

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 4 & 6 & 1 \\ -2 & 2 & 0 \end{bmatrix} \quad \text{y} \quad E_{32}E_{31}E_{21}A = U.$$

Multiplique estas matrices E para obtener una matriz M que haga la eliminación: $MA = U$.

25. Suponga que $a_{33} = 7$ y que el tercer pivote es 5. Si a_{33} se cambia a 11, el tercer pivote es _____. Si a_{33} cambia a _____, en la posición pivote hay un cero.

26. Si cada columna de A es un múltiplo de (1, 1, 1), entonces Ax siempre es un múltiplo de (1, 1, 1). Escriba un ejemplo de 3 por 3. ¿Cuántos pivotes se producen por eliminación?

27. ¿Qué matriz E_{31} resta 7 veces el renglón 1 del renglón 3? Para invertir este paso, R_{31} debe _____ 7 veces el renglón _____ al renglón _____. Multiplique E_{31} por R_{31}.

28. *a)* E_{21} resta el renglón 1 del renglón 2 y luego P_{23} intercambia los renglones 2 y 3. ¿Qué matriz $M = P_{23}E_{21}$ realiza ambos pasos a la vez?

b) P_{23} intercambia los renglones 2 y 3 y luego E_{31} resta el renglón 1 del renglón 3. ¿Qué matriz $M = E_{31}P_{23}$ realiza ambos pasos a la vez? Explique por qué las matrices M son las mismas pero las matrices E son distintas.

29. *a)* ¿Qué matriz E_{13} de 3 por 3 suma el renglón 3 al renglón 1?

b) ¿Qué matriz suma el renglón 1 al renglón 3 y *al mismo tiempo* suma el renglón 3 al renglón 1?

c) ¿Qué matriz suma el renglón 1 al renglón 3 y *luego* suma el renglón 3 al renglón 1?

30. Multiplique las siguientes matrices:

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 1 \\ 1 & 4 & 0 \end{bmatrix}.$$

31. ¿Cuáles matrices de eliminación E_{21}, E_{32} y E_{43} requieren la siguiente matriz de 4 por 4?

$$A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}.$$

Los problemas 32 a 44 son sobre la obtención y multiplicación de matrices.

32. Escriba estos problemas clásicos en forma matricial de 2 por 2 para $Ax = b$ y resuélvalos:

a) X es dos veces más viejo que Y y la suma de la edad de ambos es igual a 39.

b) $(x, y) = (2, 5)$ y $(3, 7)$ está en la recta $y = mx + c$. Encuentre m y c.

33. La parábola $y = a + bx + cx^2$ pasa por los puntos $(x, y) = (1, 4)$ y $(2, 8)$ y $(3, 14)$. Encuentre y resuelva una ecuación matricial para las incógnitas (a, b, c).

34. Multiplique las siguientes matrices en los órdenes EF, FE y E^2.

$$E = \begin{bmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & 0 & 1 \end{bmatrix} \qquad F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & c & 1 \end{bmatrix}.$$

35. *a)* Suponga que todas las columnas de B son las mismas. Entonces todas las columnas de EB son las mismas, ya que cada una es E multiplicada por _____.

b) Suponga que todos los renglones de B son [1 2 4]. Demuestre con un ejemplo que todos los renglones de EB *no* son [1 2 4]. Es cierto que todos estos renglones son _____.

36. Si E suma el renglón 1 al renglón 2, y F suma el renglón 2 al renglón 1, ¿es cierto que EF es igual a FE?

37. La primera componente de Ax es $\sum a_{1j}\, x_j = a_{11}x_1 + \cdots + a_{1n}x_n$. Escriba fórmulas para la tercera componente de Ax y el elemento (1, 1) de A^2.

38. Si $AB = I$ y $BC = I$, aplique la ley asociativa para demostrar que $A = C$.

39. A es 3 por 5, B es 5 por 3, C es 5 por 1, y D es 3 por 1. *Todos los elementos son* 1. ¿Cuáles de las siguientes operaciones matriciales son permitidas, y cuáles son los resultados?

$$BA \qquad AB \qquad ABD \qquad DBA \qquad A(B + C).$$

40. ¿Qué renglones, columnas o matrices es necesario multiplicar para encontrar,

a) la tercera columna de AB?

b) el primer renglón de AB?

c) el elemento en el renglón 3, columna 4 de AB?

d) el elemento en el renglón 1, columna 1 de CDE?

41. (Matrices de 3 por 3) Escoja la única B de modo que para toda matriz A,

a) $BA = 4A$.

b) $BA = 4B$.

c) BA tiene invertidos los renglones 1 y 3 de A y el renglón 2 sin cambio.

d) Todos los renglones de BA son los mismos que el renglón 1 de A.

42. ¿Falso o verdadero?

a) Si A^2 está definida, entonces A es necesariamente cuadrada.

b) Si AB y BA están definidas, entonces A y B son cuadradas.

c) Si AB y BA están definidas, entonces AB y BA son cuadradas.

d) Si $AB = B$, entonces $A = I$.

43. Si A es de m por n, ¿cuántas multiplicaciones por separado están implicadas si

a) A multiplica un vector x con n componentes?

b) A multiplica una matriz B de n por p? Entonces AB es m por p.

c) A se multiplica a sí misma para producir A^2? Aquí $m = n$.

44. Para demostrar que $(AB)C = A(BC)$, use los vectores columna $b_1, \cdots, b_n$ de B. Primero suponga que C tiene una sola columna c con elementos $c_1, \cdots, c_n$:

AB tiene columnas $Ab_1, \cdots, Ab_n$ y Bc tienen una columna $c_1b_1 + \cdots + c_nb_n$.

Entonces $(AB)c = c_1Ab_1 + \cdots + c_nAb_n$ es igual a $A(c_1b_1 + \cdots + c_nb_n) = A(Bc)$.

La *linealidad* proporciona la igualdad de estas dos sumas, y $(AB)c = A(Bc)$. Lo mismo es cierto para todos (todas) los (las) ____ de C. En consecuencia, $(AB)C = A(BC)$.

En los problemas 45 a 49 se utiliza multiplicación columna-renglón, y multiplicación en bloque.

45. Multiplique AB usando multiplicación de columnas por renglones:

$$AB = \begin{bmatrix} 1 & 0 \\ 2 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 3 & 0 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \begin{bmatrix} 3 & 3 & 0 \end{bmatrix} + ____ = ____.$$

46. La ***multiplicación en bloque*** separa las matrices en bloques (submatrices). Si sus formas hacen posible la multiplicación en bloque, entonces es permitida. Sustituya las siguientes x por números, y confirme que la multiplicación en bloque es exitosa.

$$\begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} AC + BD \end{bmatrix} \quad \text{y} \quad \left[\begin{array}{cc|c} x & x & x \\ x & x & x \\ \hline x & x & x \end{array}\right] \left[\begin{array}{cc|c} x & x & x \\ x & x & x \\ \hline x & x & x \end{array}\right].$$

47. Trace los cortes en A y B y AB para mostrar cómo cada una de las cuatro reglas de la multiplicación es en realidad una multiplicación en bloque para encontrar AB:

a) La matriz A multiplicada por las columnas de B.

b) Los renglones de A multiplicados por la matriz B.

c) Los renglones de A multiplicados por las columnas de B.

d) Las columnas de A multiplicadas por los renglones de B.

48. La multiplicación en bloque indica que la eliminación en la columna 1 produce

$$EA = \begin{bmatrix} 1 & 0 \\ -c/a & I \end{bmatrix} \begin{bmatrix} a & b \\ c & D \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & __ \end{bmatrix}.$$

49. *Eliminación para una matriz en bloque de 2 por 2*: Cuando $A^{-1}A = I$, el primer renglón en bloque se multiplica por CA^{-1} y se resta del segundo renglón, para encontrar el *"complemento de Schur" S*:

$$\begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & S \end{bmatrix}.$$

50. Con $i^2 = -1$, el producto $(A + iB)(x + iy)$ es $Ax + iBx + iAy - By$. Use bloques para separar la parte real de la parte imaginaria que multiplica i:

$$\begin{bmatrix} A & -B \\ ? & ? \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} Ax - By \\ ? \end{bmatrix} \begin{matrix} \text{parte real} \\ \text{parte imaginaria} \end{matrix}$$

51. Suponga que resuelve $Ax = b$ para tres miembros derechos especiales de b:

$$Ax_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{y} \quad Ax_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad Ax_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Si las soluciones x_1, x_2, x_3 son las columnas de una matriz X, ¿qué es AX?

52. Si las tres soluciones en el problema 51 son $x_1 = (1, 1, 1)$ y $x_2 = (0, 1, 1)$ y $x_3 = (0, 0, 1)$, resuelva $Ax = b$ cuando $b = (3, 5, 8)$. Pregunta de desafío: ¿Qué es A?

53. Encuentre todas las matrices

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{que satisfacen} \quad A \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} A.$$

54. Si una *matriz noroeste* A se multiplica por una *matriz sureste* B, ¿qué tipo de matrices son AB y BA? "Noroeste" y "sureste" significa ceros abajo y arriba de la antidiagonal que va de $(1, n)$ a $(n, 1)$.

55. Escriba $2x + 3y + z + 5t = 8$ como una matriz A (¿cuántos renglones tiene?) multiplicando el vector columna (x, y, z, t) para obtener b. Las soluciones llenan un plano en el espacio tetradimensional. *El plano es tridimensional sin volumen de 4-dimensiones.*

56. ¿Cuál es la matriz P_1 de 2 por 2 que proyecta el vector (x, y) sobre el eje x para producir $(x, 0)$? ¿Cuál es la matriz P_2 que proyecta el vector (x, y) sobre el eje y para producir $(0, y)$? Si se multiplica $(5, 7)$ por P_1 y luego se multiplica por P_2, se obtienen (____) y (____).

57. Escriba el producto interno de $(1, 4, 5)$ y (x, y, z) como una multiplicación de matrices Ax. A tiene un renglón. Las soluciones de $Ax = 0$ están en un ____ perpendicular al vector ____. Las columnas de A sólo están en el espacio ____dimensional.

58. En notación de MATLAB, escriba los comandos que definen la matriz A y los vectores columna x y b. ¿Cuál es el comando que prueba si $Ax = b$ o no?

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \qquad x = \begin{bmatrix} 5 \\ -2 \end{bmatrix} \qquad b = \begin{bmatrix} 1 \\ 7 \end{bmatrix}$$

59. Los comandos MATLAB A = eye(3) y v = [3:5]' producen la matriz identidad de 3 por 3 y el vector columna (3, 4, 5). ¿Cuáles son las salidas de A * v y v' * v? (¡No se requiere computadora!) Si pregunta por v * A, ¿qué ocurre?

60. Si la matriz de 4 por 4 integrada sólo por unos A = ones(4, 4) se multiplica por la columna v = ones(4, 1), ¿qué es A * v? (No se requiere computadora) Si B = eye(4) + ones(4, 4) se multiplica por w = zeros(4, 1) + 2* ones(4, 1), ¿qué es B * w?

61. Escriba una **matriz mágica** M de 3 por 3 con elementos 1, 2, . . . , 9. La suma de todos los elementos de los renglones, de las columnas y de la diagonal debe ser igual a 15. El primer renglón podría ser 8, 3, 4. ¿Cuánto es M por (1, 1, 1)? ¿A qué es igual el vector renglón [1, 1, 1] multiplicado por M?

1.5 FACTORES TRIANGULARES E INTERCAMBIOS DE RENGLONES

Se desea considerar nuevamente la eliminación para ver lo que significa en términos de matrices. El punto de partida fue el sistema modelo $Ax = b$:

$$Ax = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix} = b. \qquad (1)$$

Luego había tres pasos de eliminación, con multiplicadores 2, −1, −1:

Paso 1. Restar 2 veces la primera ecuación de la segunda.
Paso 2. Restar –1 vez la primera ecuación de la tercera.
Paso 3. Restar –1 vez la segunda ecuación de la tercera.

El resultado era un sistema equivalente $Ux = c$, con una nueva matriz de coeficientes U:

$$\textbf{Triangular superior} \qquad Ux = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 5 \\ -12 \\ 2 \end{bmatrix} = c. \qquad (2)$$

Esta matriz U es **triangular superior**: todos los elementos debajo de la diagonal son cero.

El nuevo miembro derecho c se obtuvo a partir del vector original b siguiendo los mismos pasos que llevan de A a U. La *eliminación hacia adelante* equivale a tres operaciones en renglones:

Empezar con A y b;
Aplicar los pasos 1, 2, 3 en ese orden;
Terminar con U y c.

$Ux = c$ se resuelve por sustitución hacia atrás. Aquí nos centramos en relacionar A con U.

Las matrices E para el paso 1, F para el paso 2, y G para el paso 3 se presentaron en la sección precedente. Se denominan ***matrices elementales***, y es fácil ver cómo trabajan. Para restar un múltiplo ℓ de la ecuación j de la ecuación i, *el número* $-\ell$ *se escribe en la posición* (i, j). En caso contrario, se preserva la matriz identidad, con 1s en la diagonal y 0s fuera de ésta. Luego, la multiplicación de matrices realiza la operación en los renglones.

El resultado de todos los tres pasos es $GFEA = U$. Observe que E es la primera en multiplicar a A, luego es F, y por último G. Puede multiplicarse GFE entre sí para encontrar una simple matriz que lleve A a U (y también b a c). Es una matriz triangular inferior (se omiten los ceros):

$$\textbf{De } A \textbf{ a } U \qquad GFE = \begin{bmatrix} 1 & & \\ & 1 & \\ & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ 1 & & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ -2 & 1 & \\ & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ -2 & 1 & \\ -1 & 1 & 1 \end{bmatrix}. \qquad (3)$$

Esto es bueno, pero la pregunta más importante es exactamente lo opuesto: ¿Cómo puede regresarse de U a A? ***¿Cómo pueden deshacerse los pasos de la eliminación gaussiana?***

Deshacer el paso 1 no es difícil. En lugar de restar, *se suma* dos veces el primer renglón al segundo. (¡No dos veces el segundo renglón al primero!) El resultado de realizar tanto la sustracción como la adición es recuperar la matriz identidad:

$$\begin{matrix}\textbf{La inversa de} \\ \textbf{la sustracción} \\ \textbf{es la adición}\end{matrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad (4)$$

Una operación cancela la otra. En términos matriciales, una matriz es la ***inversa*** de la otra. Si la matriz elemental E contiene al número $-\ell$ en la posición (i, j), entonces la inversa E^{-1} contiene a $+\ell$ en esa posición. Por tanto, $E^{-1}E = I$, que es la ecuación (4).

Cada paso de la eliminación puede invertirse, usando E^{-1} y F^{-1} y G^{-1}. Considero que no es malo abordar las inversas ahora, antes de la siguiente sección. El problema final es deshacer todo el proceso de una vez, y ver qué matriz regresa U a A.

Debido a que el paso 3 fue el último yendo de A a U, su matriz G debe ser la primera en ser invertida en la dirección opuesta. ¡Las inversas se presentan en orden opuesto! El segundo paso en reversa es F^{-1} y el último es E^{-1}:

$$\textbf{Regreso } \boldsymbol{a A} \textbf{ desde } U \qquad E^{-1}F^{-1}G^{-1}U = A \quad \text{es} \quad LU = A. \qquad (5)$$

Es posible sustituir $GFEA$ por U con la finalidad de ver, cómo las inversas eliminan los pasos originales.

Luego se reconoce la matriz L que regresa U a A. Se denomina L porque es *triangular inferior*. Y posee la propiedad especial de que sólo puede verse, multiplicando las tres matrices inversas en el orden correcto:

$$E^{-1}F^{-1}G^{-1} = \begin{bmatrix} 1 & & \\ \mathbf{2} & 1 & \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ \mathbf{-1} & & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ & \mathbf{-1} & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ \mathbf{2} & 1 & \\ \mathbf{-1} & \mathbf{-1} & 1 \end{bmatrix} = L. \qquad (6)$$

La cuestión especial es que ***los elementos abajo de la diagonal son los multiplicadores*** $\ell = 2, -1$, y –1. Cuando se multiplican matrices, suele no haber una forma directa de leer la respuesta. Aquí las matrices se presentan justo en el orden correcto, de modo que su producto puede escribirse de inmediato. Si la computadora almacena cada multiplicador ℓ_{ij}, el número que multiplica el renglón pivote j cuando se resta del renglón i y produce un cero en la posición i, j, entonces estos multiplicadores proporcionan un registro completo de la eliminación.

Los números ℓ_{ij} ***tienen cabida perfecta en la matriz L que lleva U de nuevo a A.***

1H ***Factorización triangular A = LU sin intercambio de renglones.*** L es triangular inferior, con 1s en la diagonal. Los multiplicadores ℓ_{ij} (tomados de la eliminación) están debajo de la diagonal. U es la matriz triangular superior que aparece después de la eliminación hacia adelante. Los elementos en la diagonal de U son los pivotes.

Ejemplo 1

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 8 \end{bmatrix} \text{ se transforma en } U = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \text{ con } L = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}. \text{ Luego, } LU = A.$$

Ejemplo 2 Lo cual necesita un intercambio de renglones

$$A = \begin{bmatrix} 0 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{no puede factorizarse en} \quad A = LU.$$

Ejemplo 3 (con todos los pivotes y multiplicadores iguales a 1)

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = LU.$$

De A a U hay restas de renglones. De U a A hay sumas de renglones.

Ejemplo 4 (cuando U es la identidad y L es la misma que A)

$$\textbf{Caso triangular inferior} \quad A = \begin{bmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{bmatrix}.$$

Los pasos de eliminación en esta A son fáciles: i) E resta ℓ_{21} veces el renglón 1 del renglón 2, ii) F resta ℓ_{31} veces el renglón 1 del renglón 3, y iii) G resta ℓ_{32} veces el renglón 2 del renglón 3. El resultado es la matriz identidad $U = I$. Las inversas de E, F, y G producen de regreso A:

E^{-1} aplicada a F^{-1} aplicada a G^{-1} aplicada a I produce A.

$$\begin{bmatrix} 1 & & \\ \ell_{21} & 1 & \\ & & 1 \end{bmatrix} \begin{matrix}\text{Multi-}\\ \text{plicada}\\ \text{por}\end{matrix} \begin{bmatrix} 1 & & \\ & 1 & \\ \ell_{31} & & 1 \end{bmatrix} \begin{matrix}\text{Multi-}\\ \text{plicada}\\ \text{por}\end{matrix} \begin{bmatrix} 1 & & \\ & 1 & \\ & \ell_{32} & 1 \end{bmatrix} \begin{matrix}\text{es}\\ \text{igual}\\ \text{a}\end{matrix} \begin{bmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{bmatrix}.$$

El orden es correcto para que los ℓs estén en su posición. ¡Esto ocurre siempre! Observe que los paréntesis en $E^{-1}F^{-1}G^{-1}$ no fueron necesarios debido a la ley asociativa.

$A = LU$: El caso n por n

La factorización $A = LU$ es tan importante que es necesario comentar más acerca de ella. Normalmente se la ignoraba en cursos de álgebra lineal cuando la orientación de éstos era la parte abstracta. O quizá se pensaba que era demasiado difícil, pero es necesario abordarla. Si el último ejemplo 4 permite cualquier U en vez de la U particular $U = I$, es posible percatarse de cómo trabaja la regla en general. ***La matriz L, aplicada a U, regresa A***:

$$A = LU \qquad \begin{bmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{bmatrix} \begin{bmatrix} \text{renglón 1 de } U \\ \text{renglón 2 de } U \\ \text{renglón 3 de } U \end{bmatrix} = A \text{ original.} \tag{7}$$

La demostración es *aplicar los pasos de eliminación*. En el miembro derecho, éstos llevan A a U. En el miembro izquierdo reducen L a I, como en el ejemplo 4. (El primer paso resta ℓ_{21} veces $(1, 0, 0)$ del segundo renglón, con lo cual se elimina ℓ_{21}.) Ambos miembros de (7) terminan iguales a la misma matriz U, y todos los pasos para llegar ahí son reversibles. En consecuencia (7) es correcta y $A = LU$.

$A = LU$ es tan crucial, y tan bella, que el problema 8 al final de esta sección sugiere un segundo método. Aquí se están escribiendo matrices de 3 por 3, aunque el lector puede darse cuenta de la forma en que los razonamientos se aplican a matrices más grandes. A continuación se proporciona otro ejemplo, y luego se empieza a utilizar $A = LU$.

Ejemplo 5 ($A = LU$, con ceros en los espacios vacíos)

$$A = \begin{bmatrix} 1 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & -1 & 1 & \\ & & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & & \\ & 1 & -1 & \\ & & 1 & -1 \\ & & & 1 \end{bmatrix}.$$

Esto demuestra cómo una matriz A con tres diagonales tiene factores L y U con dos diagonales. Este ejemplo proviene de un problema importante en ecuaciones diferenciales (véase la sección 1.7). La segunda diferencia en A es la diferencia hacia atrás L multiplicada por una diferencia hacia adelante U.

Un sistema lineal = Dos sistemas triangulares

Hay un punto práctico importante que mencionar sobre $A = LU$. Es más que un simple registro de pasos de eliminación; L y U son las matrices derechas para resolver $Ax = b$. De hecho, ¡es posible descartar a A! De b a c se va por eliminación directa (que utiliza L), y de c a x se va por sustitución hacia atrás (que usa U). Esto puede y debe hacerse sin A:

$$\textbf{Separación de } Ax = b \qquad \textit{Primero} \quad Lc = b \quad \textit{y luego} \quad Ux = c. \tag{8}$$

La segunda ecuación se multiplica por L para obtener $LUx = Lc$, que es $Ax = b$. Cada sistema triangular se resuelve rápidamente. Esto es exactamente lo que un buen código de eliminación hace:

1. ***Factoriza*** (a partir de A encuentran sus factores L y U).
2. ***Resuelve*** (a partir de L y U y b encuentra la solución x).

La separación entre ***Factoriza*** y ***Resuelve*** significa que es posible procesar una serie de bs. La subrutina ***Resuelve*** obedece la ecuación (8): dos sistemas triangulares en $n^2/2$ pasos cada uno. **La solución para cualquier miembro derecho nuevo b puede encontrarse en sólo n^2 operaciones.** Esto es mucho menos que los $n^3/3$ pasos necesarios para factorizar A en el miembro izquierdo.

Ejemplo 6 Esta es la matriz previa A con miembro derecho $b = (1, 1, 1, 1)$.

$$Ax = b \quad \begin{aligned} x_1 - x_2 &= 1 \\ -x_1 + 2x_2 - x_3 &= 1 \\ -x_2 + 2x_3 - x_4 &= 1 \\ -x_3 + 2x_4 &= 1 \end{aligned} \quad \text{se separa en } Lc = b \text{ y } Ux = c.$$

$$Lc = b \quad \begin{aligned} c_1 &= 1 \\ -c_1 + c_2 &= 1 \\ -c_2 + c_3 &= 1 \\ -c_3 + c_4 &= 1 \end{aligned} \quad \text{proporciona } c = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}.$$

$$Ux = c \quad \begin{aligned} x_1 - x_2 &= 1 \\ x_2 - x_3 &= 2 \\ x_3 - x_4 &= 3 \\ x_4 &= 4 \end{aligned} \quad \text{proporciona } x = \begin{bmatrix} 10 \\ 9 \\ 7 \\ 4 \end{bmatrix}.$$

Para estas "matrices tridiagonales" especiales, la operación cae desde n^2 hasta $2n$. Usted verá cómo $Lc = b$ se resuelve *hacia adelante* (c_1 viene antes de c_2). Esto es precisamente lo que ocurre durante la eliminación hacia adelante. Luego, $Ux = c$ se resuelve hacia atrás (x_4 antes de x_3).

Observación 1 La forma LU es "no simétrica" en la diagonal: L tiene 1s donde U tiene los pivotes. Esto es fácil de corregir: ***U* se divide entre una matriz pivote diagonal *D*:**

$$\textbf{Factorizar } D \qquad U = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \cdot & \\ & & & d_n \end{bmatrix} \begin{bmatrix} 1 & u_{12}/d_1 & u_{13}/d_1 & \cdot \\ & 1 & u_{23}/d_2 & \cdot \\ & & \cdot & \cdot \\ & & & 1 \end{bmatrix}. \qquad (9)$$

En el último ejemplo, todos los pivotes eran $d_i = 1$. En ese caso $D = I$. Pero esto fue muy excepcional, y normalmente LU es diferente de LDU (lo cual también se escribe como LDV).

> ***La factorización triangular puede escribirse como A = LDU, donde L y U tienen 1s. en la diagonal, y D es la matriz diagonal de pivotes.***

Siempre que se observa LDU o LDV, se entiende que U o V tiene 1s en la diagonal: cada renglón se dividió entre el pivote en D. Luego, L y U se tratan de la misma forma. A continuación se muestra un ejemplo de separación de LU en LDU.

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ & -2 \end{bmatrix} = \begin{bmatrix} 1 & \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & \\ & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ & 1 \end{bmatrix} = LDU.$$

Lo anterior tiene los 1s en las diagonales de L y U, y los pivotes 1 y -2 en D.

Observación 2 Al describir cada paso de eliminación, quizá dimos la impresión de que los cálculos deben realizarse en ese orden. Esto es erróneo. Hay *algo* de libertad, y hay un "algoritmo Crout" que arregla los cálculos de una manera ligeramente distinta.

En las L, D, y U finales no hay libertad. Esta es la cuestión más importante:

> **1I** Si $A = L_1D_1U_1$ y también $A = L_2D_2U_2$, donde las L son triangulares inferiores con diagonales unitarias, las U son triangulares superiores con diagonales unitarias, y las D son matrices diagonales sin ceros en la diagonal, entonces $L_1 = L_2$, $D_1 = D_2$, $U_1 = U_2$. La factorización LDU y la factorización LU están determinadas de manera única por A.

La demostración constituye un buen ejercicio en la siguiente sección.

Intercambios de renglones, y matrices de permutación

Ahora es necesario enfrentar un problema que se ha evitado hasta el momento: el número que se espera utilizar como pivote puede ser cero. Esto puede ocurrir en medio de un cálculo. Ocurrirá casi al principio si $a_{11} = 0$. Un simple ejemplo es

$$\textbf{Cero en la posición pivote} \qquad \begin{bmatrix} \mathbf{0} & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

La dificultad es evidente; ningún múltiplo de la primera ecuación elimina al coeficiente 3.

El remedio es igualmente claro. ***Intercambiar las dos ecuaciones***, moviendo el elemento 3 hacia el pivote. En este ejemplo, la matriz pudo transformarse en triangular superior:

$$\textbf{Intercambio de renglones} \qquad \begin{aligned} 3u + 4v &= b_2 \\ 2v &= b_1 \end{aligned}$$

Para expresar esto en términos matriciales, se requiere la ***matriz de permutación*** P que produce el intercambio de renglones. Se obtiene al intercambiar los renglones de I:

$$\textbf{Permutación} \qquad P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{y} \quad PA = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 0 & 2 \end{bmatrix}.$$

P tiene el mismo efecto sobre b, intercambiando b_1 y b_2. *El nuevo sistema es* $PAx = Pb$. Las incógnitas u y v *no* se invierten en un intercambio de renglones.

Una matriz de permutación P tiene los mismos renglones que la identidad (en algún orden). En cada renglón y columna hay un simple "1". La matriz de permutación más común es $P = I$ (no intercambia nada). El producto de dos matrices de permutación es otra permutación: los renglones de I se reordenan dos veces.

Después de $P = I$, las permutaciones más sencillas intercambian dos renglones. Otras permutaciones intercambian más renglones. **Hay $n! = (n)(n-1)\cdots(1)$ permutaciones de tamaño n.** El renglón 1 tiene n opciones; luego, el renglón 2 tiene $n - 1$ opciones, y finalmente el último renglón sólo tiene una opción. Es posible mostrar todas las permutaciones de 3 por 3 (hay $3! = (3)(2)(1) = 6$ matrices):

$$I = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \quad P_{21} = \begin{bmatrix} & 1 & \\ 1 & & \\ & & 1 \end{bmatrix} \quad P_{32}P_{21} = \begin{bmatrix} & 1 & \\ & & 1 \\ 1 & & \end{bmatrix}$$

$$P_{31} = \begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix} \quad P_{32} = \begin{bmatrix} 1 & & \\ & & 1 \\ & 1 & \end{bmatrix} \quad P_{21}P_{32} = \begin{bmatrix} & & 1 \\ 1 & & \\ & 1 & \end{bmatrix}.$$

Hay 24 matrices de permutación de orden $n = 4$. Sólo hay dos matrices de permutación de orden $n = 2$; a saber,

$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \quad \text{y} \quad \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}.$$

Una vez que se conoce sobre inversas y traspuestas (en la siguiente sección se definen A^{-1} y A^{T}), se descubre un hecho importante: ***P^{-1} siempre es igual a P^{T}***.

Un cero en la ubicación del pivote origina dos posibilidades: ***El problema puede ser fácil de arreglar, o puede ser grave.*** Esto se decide al observar *abajo del cero*. Si más abajo en la misma columna hay un elemento distinto de cero, entonces se realiza un intercambio de renglones. El elemento distinto de cero se convierte en el pivote necesario, y la eliminación puede continuar de nuevo:

$$A = \begin{bmatrix}0 & a & b\\0 & 0 & c\\d & e & f\end{bmatrix} \qquad \begin{matrix}d = 0 & \Longrightarrow & \text{ningún primer pivote}\\ a = 0 & \Longrightarrow & \text{ningún segundo pivote}\\ c = 0 & \Longrightarrow & \text{ningún tercer pivote}\end{matrix}$$

Si $d = 0$, entonces el problema es incurable y esta matriz es ***singular***. No hay esperanza de una solución única de $Ax = b$. Si d *no* es cero, un intercambio P_{13} de los renglones 1 y 3 mueve d al pivote. No obstante, la siguiente posición pivote también contiene un cero. El número a está ahora abajo del cero (el e arriba del cero es inútil). Si a no es cero, entonces se requiere otro intercambio de renglones P_{23}:

$$P_{13} = \begin{bmatrix}0 & 0 & 1\\0 & 1 & 0\\1 & 0 & 0\end{bmatrix} \quad \text{y} \quad P_{23} = \begin{bmatrix}1 & 0 & 0\\0 & 0 & 1\\0 & 1 & 0\end{bmatrix} \quad \text{y} \quad P_{23}P_{13}A = \begin{bmatrix}\boldsymbol{d} & e & f\\0 & \boldsymbol{a} & b\\0 & 0 & \boldsymbol{c}\end{bmatrix}$$

Una cuestión más: La permutación $P_{23}P_{13}$ realiza los dos intercambios de renglones a la vez:

$$\textbf{P_{13} actúa primero} \qquad P_{23}P_{13} = \begin{bmatrix}1 & 0 & 0\\0 & 0 & 1\\0 & 1 & 0\end{bmatrix}\begin{bmatrix}0 & 0 & 1\\0 & 1 & 0\\1 & 0 & 0\end{bmatrix} = \begin{bmatrix}0 & 0 & 1\\1 & 0 & 0\\0 & 1 & 0\end{bmatrix} = P.$$

De haberlo sabido, hubiera sido posible multiplicar A por P en primer lugar. Con los renglones en el orden correcto PA, cualquier matriz no singular está lista para el proceso de eliminación.

Eliminación en pocas palabras: $PA = LU$

La cuestión más importante es la siguiente: Si la eliminación puede completarse con ayuda de intercambios de renglones, entonces es posible suponer que estos intercambios se realizan primero (por P). *La matriz PA no requiere intercambios de renglones*. En otras palabras, PA permite la factorización normal en L multiplicada por U. La teoría de la eliminación gaussiana puede resumirse en unas cuantas líneas:

1J En el caso ***no singular***, hay una matriz de permutación P que reordena los renglones de A para evitar la presencia de ceros en las posiciones pivote. Así, $Ax = b$ tiene una *solución única*:

Con los renglones reordenados de antemano, PA puede factorizarse en LU.

En el caso ***singular***, ninguna P puede producir un conjunto completo de pivotes: la eliminación fracasa.

En la práctica, también es posible considerar un intercambio de renglones cuando el pivote original está *próximo* a cero: aun cuando no sea exactamente igual a cero. Cuando se elige un pivote más grande se reduce el error por redondeo.

Debe tenerse cuidado con L. Suponga que la eliminación resta el renglón 1 del renglón 2, creando $\ell_{21} = 1$. Luego suponga que intercambia los renglones 2 y 3. Si este intercambio se realiza de antemano, el multiplicador cambia a $\ell_{31} = 1$ en $PA = LU$.

Ejemplo 7

$$A = \begin{bmatrix}1 & 1 & 1\\1 & 1 & 3\\2 & 5 & 8\end{bmatrix} \to \begin{bmatrix}1 & 1 & 1\\0 & \boldsymbol{0} & 2\\0 & 3 & 6\end{bmatrix} \to \begin{bmatrix}1 & 1 & 1\\0 & 3 & 6\\0 & 0 & 2\end{bmatrix} = U. \tag{10}$$

Ese intercambio de renglones recupera LU, pero no $\ell_{31} = 1$ y $\ell_{21} = 2$:

$$P = \begin{bmatrix}1 & 0 & 0\\0 & 0 & 1\\0 & 1 & 0\end{bmatrix} \quad \text{y} \quad L = \begin{bmatrix}1 & 0 & 0\\2 & 1 & 0\\1 & 0 & 1\end{bmatrix} \quad \text{y} \quad PA = LU. \tag{11}$$

En MATLAB , A([r k],:) intercambia el renglón k con el renglón r abajo del renglón k (donde se ha encontrado el k-ésimo pivote). Las matrices L y P se actualizan de la misma manera. Al principio, $P = I$ y signo $= +1$:

```
A([r k], : ) = A([k r], : );
L([r k],1:k−1) = L([k r],1:k−1);
P([r k], : ) = P([k r], : );
sign = −sign
```

El "**signo**" de P indica si el número de intercambios de renglones es par (signo $= +1$) o impar (signo $= -1$). Un intercambio de renglones invierte el signo. El valor final del signo es el **determinante de P** y no depende del orden de los intercambios de renglones.

En resumen: Un buen código de eliminación ahorra L y U y P. Estas matrices contienen la información que originalmente estaba en A, y la contienen en una forma más utilizable. $Ax = b$ se reduce a dos sistemas triangulares. Este es el equivalente práctico del cálculo que se presenta a continuación: *encontrar la matriz inversa A^{-1} y la solución $x = A^{-1}b$.*

Conjunto de problemas 1.5

1. ¿Cuándo es no singular una matriz triangular superior (un conjunto completo de pivotes)?
2. ¿Qué múltiplo ℓ_{32} del renglón 2 de A resta la eliminación del renglón 3 de A? Use la forma factorizada

$$A = \begin{bmatrix}1 & 0 & 0\\2 & 1 & 0\\1 & 4 & 1\end{bmatrix}\begin{bmatrix}5 & 7 & 8\\0 & 2 & 3\\0 & 0 & 6\end{bmatrix}.$$

¿Cuáles son los pivotes? ¿Se requiere un intercambio de renglones?

3. Multiplique la matriz $L = E^{-1}F^{-1}G^{-1}$ en la ecuación (6) por GFE en la ecuación (3):

$$\begin{bmatrix}1 & 0 & 0\\2 & 1 & 0\\-1 & -1 & 1\end{bmatrix} \text{ multiplicada por } \begin{bmatrix}1 & 0 & 0\\-2 & 1 & 0\\-1 & 1 & 1\end{bmatrix}.$$

También multiplique en orden opuesto. *¿Por qué las respuestas son iguales?*

4. Aplique eliminación para obtener los factores L y U de

$$A = \begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & 4 \\ 1 & 4 & 8 \end{bmatrix}.$$

5. Factorice A en LU, y escriba el sistema triangular superior $Ux = c$ que aparece después de la eliminación, para

$$Ax = \begin{bmatrix} 2 & 3 & 3 \\ 0 & 5 & 7 \\ 6 & 9 & 8 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 5 \end{bmatrix}.$$

6. Encuentre E^2 y E^8 y E^{-1} si

$$E = \begin{bmatrix} 1 & 0 \\ 6 & 1 \end{bmatrix}.$$

7. Encuentre los productos FGH y HGF si (donde se han omitido los ceros triangulares superiores)

$$F = \begin{bmatrix} 1 & & & \\ 2 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad G = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 2 & 1 & \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad H = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & 2 & 1 \end{bmatrix}.$$

8. (***Segunda demostración de*** $A = LU$) El tercer renglón de U proviene del tercer renglón de A al restar múltiplos de los renglones 1 y 2 (¡*de* U!):

renglón 3 de U = renglón 3 de $A - \ell_{31}$(renglón 1 de U) $- \ell_{32}$(renglón 2 de U).

a) ¿Por qué se restan renglones de U y no de A? Respuesta: Debido a que en ese momento se utiliza un renglón pivote ______.

b) La ecuación anterior es la misma que

renglón 3 de $A = \ell_{31}$(renglón 1 de U) $- \ell_{32}$(renglón 2 de U) + 1(renglón 3 de U).

¿Qué regla de la multiplicación de matrices multiplica por U este renglón 3 de L?

Los otros renglones de LU coinciden de manera semejante con los renglones de A.

9. *a*) ¿En qué condiciones el siguiente producto es no singular?

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} d_1 & & \\ & d_2 & \\ & & d_3 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

b) Resuelva el sistema $Ax = b$ empezando con $Lc = b$:

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = b.$$

10. *a*) ¿Por qué para resolver cada uno de los sistemas $Lc = b$ y $Ux = c$ se requieren aproximadamente $n^2/2$ pasos de multiplicación-sustracción?

b) ¿Cuántos pasos requiere la eliminación para resolver 10 sistemas con la misma matriz de coeficientes A de 60 por 60?

11. Resuelva lo siguiente como dos sistemas triangulares, sin multiplicar LU para encontrar A:

$$LUx = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}.$$

12. ¿Cómo es posible factorizar A en un producto UL, triangular superior multiplicado por un sistema triangular inferior? ¿Tendrían los mismos factores que en $A = LU$?

13. Resuelva lo siguiente por eliminación, intercambiando renglones cuando sea necesario:

$$\begin{aligned} u + 4v + 2w &= -2 \\ -2u - 8v + 3w &= 32 \\ v + w &= 1 \end{aligned} \quad \text{y} \quad \begin{aligned} v + w &= 0 \\ u + v \quad &= 0 \\ u + v + w &= 1. \end{aligned}$$

¿Qué matrices de permutación se requieren?

14. Escriba todas las matrices de permutación de 3 por 3, incluyendo $P = I$. Identifique sus inversas, que también son matrices de permutación. Las inversas satisfacen $PP^{-1} = I$, y están en la misma lista.

15. Encuentre (y compruebe) las factorizaciones $PA = LDU$ para

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 1 \end{bmatrix}.$$

16. Encuentre una matriz de permutación de 4 por 4 que requiera tres intercambios de renglones para llegar al final de la eliminación (que es $U = I$).

17. La forma menos conocida $A = LPU$ intercambia renglones sólo al final:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 3 \\ 2 & 5 & 8 \end{bmatrix} \rightarrow L^{-1}A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 3 & 6 \end{bmatrix} = PU = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & 6 \\ 0 & 0 & 2 \end{bmatrix}.$$

¿Cuál es L en este caso? Al comparar con $PA = LU$ en el recuadro 1J, ahora los multiplicadores permanecen en su sitio (ℓ_{21} es 1 y ℓ_{31} es 2 cuando $A = LPU$).

18. Decida si los siguientes sistemas son singulares o no singulares, y si no tienen solución, una solución o una infinidad de soluciones:

$$\begin{aligned} v - w &= 2 \\ u - v \quad &= 2 \\ u \quad - w &= 2 \end{aligned} \quad \text{y} \quad \begin{aligned} v - w &= 0 \\ u - v \quad &= 0 \\ u \quad - w &= 0 \end{aligned} \quad \text{y} \quad \begin{aligned} v + w &= 1 \\ u + v \quad &= 1 \\ u \quad + w &= 1. \end{aligned}$$

19. ¿Cuáles números a, b, c conducen a intercambios de renglones? ¿Cuáles hacen singular a la matriz?

$$A = \begin{bmatrix} 1 & 2 & 0 \\ a & 8 & 3 \\ 0 & b & 5 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} c & 2 \\ 6 & 4 \end{bmatrix}.$$

En los problemas 20 a 31 se calcula la factorización $A = LU$ (y también $A = LDU$).

20. La eliminación hacia adelante cambia $\begin{bmatrix}1 & 1\\1 & 2\end{bmatrix}x = b$ a una triangular $\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}x = c$:

$$\begin{matrix}x + y = 5\\ x + 2y = 7\end{matrix} \longrightarrow \begin{matrix}x + y = 5\\ y = 2\end{matrix} \qquad \begin{bmatrix}1 & 1 & \mathbf{5}\\1 & 2 & \mathbf{7}\end{bmatrix} \longrightarrow \begin{bmatrix}1 & 1 & \mathbf{5}\\0 & 1 & \mathbf{2}\end{bmatrix}.$$

Ese paso restó $\ell_{21} =$ _____ veces el renglón 1 del renglón 2. El paso inverso *suma* ℓ_{21} veces el renglón 1 al renglón 2. La matriz para ese paso es $L =$ _____. Al multiplicar esta L por el sistema triangular $\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}x = \begin{bmatrix}5\\2\end{bmatrix}$ se obtiene _____ = _____. En otras palabras, L multiplica a $Ux = c$ para obtener _____.

21. (Cambio a 3 por 3) La eliminación hacia adelante cambia $Ax = b$ a un sistema triangular $Ux = c$:

$$\begin{matrix}x + y + z = 5\\ x + 2y + 3z = 7\\ x + 3y + 6z = 11\end{matrix} \qquad \begin{matrix}x + y + z = 5\\ y + 2z = 2\\ 2y + 5z = 6\end{matrix} \qquad \begin{matrix}x + y + z = 5\\ y + 2z = 2\\ z = 2.\end{matrix}$$

La ecuación $z = 2$ en $Ux = c$ proviene de la ecuación original $x + 3y + 6z = 11$ en $Ax = b$ al restar $\ell_{31} =$ _____ veces la ecuación 1 y $\ell_{32} =$ _____ veces la ecuación *final* 2. Lo anterior se invierte para recuperar $[1\ 3\ 6\ 11]$ en $[A\ \ b]$ de la $[1\ 1\ 1\ 5]$ final y $[0\ 1\ 2\ 2]$ y $[0\ 0\ 1\ 2]$ en $[U\ \ c]$:

Renglón 3 de $[A\ \ b] = (\ell_{31}$ renglón 1 + ℓ_{32} renglón 2 + 1 renglón 3) de $[U\ \ c]$.

En notación de matrices, esto es una multiplicación por L. Así, $A = LU$ y $b = Lc$.

22. ¿Cuáles son los sistemas triangulares de 3 por 3 $Lc = b$ y $Ux = c$ del problema 21? Compruebe que $c = (5, 2, 2)$ resuelve el primero. ¿Cuál es el vector x que resuelve el segundo?

23. ¿Cuáles son las dos matrices de eliminación E_{21} y E_{32} que transforman la matriz A en la forma triangular superior $E_{32}E_{21}A = U$? Multiplique por E_{32}^{-1} y E_{21}^{-1} para factorizar A en LU $E_{21}^{-1}E_{32}^{-1}U$:

$$A = \begin{bmatrix}1 & 1 & 1\\2 & 4 & 5\\0 & 4 & 0\end{bmatrix}.$$

24. ¿Cuáles son las tres matrices de eliminación E_{21}, E_{31} y E_{32} que transforman la matriz A en la forma triangular superior $E_{32}E_{31}E_{21}A = U$? Multiplique por E_{32}^{-1}, E_{31}^{-1} y E_{21}^{-1} para factorizar A en LU, donde $L = E_{21}^{-1}E_{31}^{-1}E_{32}^{-1}$. Encuentre L y U:

$$A = \begin{bmatrix}1 & 0 & 1\\2 & 2 & 2\\3 & 4 & 5\end{bmatrix}.$$

25. Cuando en la posición pivote aparece un cero, ¡$A = LU$ *no es posible*! (Se requieren pivotes diferentes de cero d, f, i en U.) Demuestre directamente por qué los dos casos siguientes son imposibles:

$$\begin{bmatrix}0 & 1\\2 & 3\end{bmatrix} = \begin{bmatrix}1 & 0\\ \ell & 1\end{bmatrix}\begin{bmatrix}d & e\\0 & f\end{bmatrix} \qquad \begin{bmatrix}1 & 1 & 0\\1 & 1 & 2\\1 & 2 & 1\end{bmatrix} = \begin{bmatrix}1 & & \\ \ell & 1 & \\ m & n & 1\end{bmatrix}\begin{bmatrix}d & e & g\\ & f & h\\ & & i\end{bmatrix}.$$

26. ¿Qué número c produce un cero en la segunda posición pivote? Se requiere un intercambio de renglones y $A = LU$ no es posible. ¿Qué número c produce un cero en la tercera posición pivote? Así, un intercambio de renglones no puede ayudar y la eliminación falla:

$$A = \begin{bmatrix}1 & c & 0\\2 & 4 & 1\\3 & 5 & 1\end{bmatrix}.$$

27. ¿Cuáles son L y D para esta matriz A? ¿Cuál es U en $A = LU$ y cuál es la nueva U en $A = LDU$?

$$A = \begin{bmatrix}2 & 4 & 8\\0 & 3 & 9\\0 & 0 & 7\end{bmatrix}.$$

28. A y B son simétricas a través de la diagonal (porque $4 = 4$). Encuentre sus factorizaciones triples LDU, y escriba cómo está relacionada U con L para estas matrices simétricas:

$$A = \begin{bmatrix}2 & 4\\4 & 11\end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix}1 & 4 & 0\\4 & 12 & 4\\0 & 4 & 0\end{bmatrix}.$$

29. (Recomendado) Calcule L y U para la matriz simétrica

$$A = \begin{bmatrix}a & a & a & a\\a & b & b & b\\a & b & c & c\\a & b & c & d\end{bmatrix}.$$

Encuentre cuatro condiciones sobre a, b, c, d para obtener $A = LU$ con cuatro pivotes.

30. Encuentre L y U para la matriz no simétrica:

$$A = \begin{bmatrix}a & r & r & r\\a & b & s & s\\a & b & c & t\\a & b & c & d\end{bmatrix}.$$

Encuentre cuatro condiciones sobre a, b, c, d, r, s, t para obtener $A = LU$ con cuatro pivotes.

31. Las *matrices tridiagonales* tienen elementos iguales a cero, excepto en la diagonal principal y en las dos diagonales adyacentes. Factorice lo siguiente en $A = LU$ y $A = LDV$:

$$A = \begin{bmatrix}1 & 1 & 0\\1 & 2 & 1\\0 & 1 & 2\end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix}a & a & 0\\a & a+b & b\\0 & b & b+c\end{bmatrix}.$$

32. Resuelva el sistema triangular $Lc = b$ para encontrar c. Luego, resuelva $Ux = c$ para encontrar x:

$$L = \begin{bmatrix}1 & 0\\4 & 1\end{bmatrix} \quad \text{y} \quad U = \begin{bmatrix}2 & 4\\0 & 1\end{bmatrix} \quad \text{y} \quad b = \begin{bmatrix}2\\11\end{bmatrix}.$$

Por seguridad, encuentre $A = LU$ y resuelva $Ax = b$ como de costumbre. Identifique c cuando lo vea.

33. Resuelva $Lc = b$ para encontrar c. Luego resuelva $Ux = c$ para encontrar x. ¿Cuál era la matriz A?

$$L = \begin{bmatrix}1 & 0 & 0\\1 & 1 & 0\\1 & 1 & 1\end{bmatrix} \quad \text{y} \quad U = \begin{bmatrix}1 & 1 & 1\\0 & 1 & 1\\0 & 0 & 1\end{bmatrix} \quad \text{y} \quad b = \begin{bmatrix}4\\5\\6\end{bmatrix}.$$

34. Si A y B tienen elementos distintos de cero en las posiciones indicadas con x, ¿qué ceros siguen siendo ceros en sus factores L y U?

$$A = \begin{bmatrix} x & x & x & x \\ x & x & x & 0 \\ 0 & x & x & x \\ 0 & 0 & x & x \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} x & x & x & 0 \\ x & x & 0 & x \\ x & 0 & x & x \\ 0 & x & x & x \end{bmatrix}.$$

35. (Importante) Si A tiene pivotes 2, 7, 6 sin intercambios de renglones, ¿cuáles son los pivotes para la parte superior izquierda de la submatriz B de 2 por 2 (sin el renglón 3 y la columna 3)? Explique su respuesta.

36. Empezando con una matriz A de 3 por 3 con pivotes 2, 7, 6, añada un cuarto renglón y una cuarta columna para obtener M. ¿Cuáles son los tres primeros pivotes de M, y por qué? ¿Cuáles son el cuarto renglón y la cuarta columna que aseguran la obtención de un 9 en el cuarto pivote?

37. Use chol(pascal(5)) para encontrar los factores triangulares de pascal(5) de MATLAB. ¡Intercambios de renglones en $[L, U]$ = lu(pascal(5)) arruinan el patrón de Pascal!

38. (Repaso) ¿Para qué números c es imposible $A = LU$ con tres pivotes?

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & c & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

39. Calcule la diferencia temporal para cada nuevo miembro derecho b cuando $n = 800$. Establezca A = rand(800), b = rand(800,1) y B = rand(800,9). Compare los tiempos de tic; A\b; toc y tic A\B; toc (que resuelve para 9 miembros derechos).

Los problemas 40 a 48 son sobre matrices de permutación.

40. Hay 12 permutaciones "*pares*" de (1, 2, 3, 4) con un *número par de intercambios*. Dos de éstas son (1, 2, 3, 4) sin intercambios y (4, 3, 2, 1) con dos intercambios. Enumere los otros 10. En vez de escribir cada matriz de 4 por 4, use los números 4, 3, 2, 1 para proporcionar la posición del 1 en cada renglón.

41. ¿Con cuántos intercambios se permuta (5, 4, 3, 2, 1) de vuelta a (1, 2, 3, 4, 5)? ¿Cuántos intercambios son necesarios para ir de (6, 5, 4, 3, 2, 1) a (1, 2, 3, 4, 5, 6)? Uno es par y el otro es impar. Para ir de $(n, \ldots, 1)$ a $(1, \ldots, n)$, demuestre que $n = 100$ y 101 son pares, y que $n = 102$ y 103 son impares.

42. Si P_1 y P_2 son matrices de permutación, también lo es P_1P_2. Esta última matriz sigue teniendo los renglones de I en algún orden. Proporcione ejemplos con $P_1P_2 \neq P_2P_1$ y $P_3P_4 = P_4P_3$.

43. (Intente responder esta pregunta). ¿Cuál es la permutación que hace triangular superior a PA? ¿Cuál es la permutación que hace triangular inferior a P_1AP_2? ***Al multiplicar A por la derecha por P_2 se intercambia el _____ de A.***

$$A = \begin{bmatrix} 0 & 0 & 6 \\ 1 & 2 & 3 \\ 0 & 4 & 5 \end{bmatrix}.$$

44. Encuentre una matriz de permutación de 3 por 3 con $P^3 = I$ (pero no $P = I$). Encuentre una permutación de 4 por 4 $\widehat{P}$ con $\widehat{P}^4 \neq I$.

45. Si se toman potencias de una permutación, ¿por qué alguna P^k termina siendo igual a I? Encuentre una permutación de 5 por 5 P de modo que la menor potencia que es igual a I es P^6. (Éste es un problema de desafío. Combine un bloque de 2 por 2 con un bloque de 3 por 3.)

46. La matriz P que multiplica a (x, y, z) para obtener (z, x, y) también es una matriz de rotación. Encuentre P y P^3. El eje de rotación $a = (1, 1, 1)$ no se mueve, y es igual a Pa. ¿Cuál es el ángulo de rotación de $v = (2, 3, -5)$ a $Pv = (-5, 2, 3)$?

47. Si P es cualquier matriz de permutación, encuentre un vector x diferente de cero de modo que $(I - P)x = 0$. (Esto significa que $I - P$ no tiene inversa, y que su determinante es cero).

48. Si P tiene 1s en su antidiagonal desde $(1, n)$ hasta $(n, 1)$, describa PAP.

1.6 INVERSAS Y TRASPUESTAS

La inversa de una matriz de n por n es otra matriz de n por n. La inversa de A se denota por A^{-1} (lo cual se lee "A inversa"). La propiedad fundamental es sencilla: *Si se multiplica por A y luego se multiplica por A^{-1}, se regresa a la posición inicial*:

Matriz inversa Si $b = Ax$, entonces $A^{-1}b = x$.

Así, $A^{-1}Ax = x$. La matriz A^{-1} multiplicada por la matriz A es la matriz identidad. ***No todas las matrices tienen inversas. Una inversa es imposible cuando Ax es cero y x es diferente de cero.*** Por tanto, A^{-1} debería tener que regresar de $Ax = 0$ a x. Ninguna matriz puede multiplicar ese vector cero Ax y producir un vector x diferente de cero.

Los objetivos son definir la matriz inversa, calcularla y aplicarla, cuando A^{-1} existe, y luego comprender cuáles matrices no tienen inversas.

1K La inversa de A es una matriz B tal que $BA = I$ y $AB = I$. Hay cuando mucho una B, y ésta se denota por A^{-1}:

$$A^{-1}A = I \quad \text{y} \quad AA^{-1} = I \tag{1}$$

Nota 1 ***La inversa existe si y sólo si la eliminación produce n pivotes*** (se permiten intercambios de renglones). La eliminación resuelve $Ax = b$ sin encontrar explícitamente A^{-1}.

Nota 2 La matriz A no puede tener dos inversas distintas. Suponga que $BA = I$ y también que $AC = I$. Entonces $B = C$, según la siguiente "demostración por paréntesis":

$$B(AC) = (BA)C \text{ proporciona } BI = IC, \text{ que es } B = C. \tag{2}$$

Esto demuestra que una *inversa izquierda* B (que multiplica por la izquierda) y una *inversa derecha* C (que multiplica a A por la derecha para obtener $AC = I$) deben ser la *misma matriz*.

Nota 3 Si A es invertible, la sola y única solución de $Ax = b$ es $x = A^{-1}b$:

Si $Ax = b$ se multiplica por A^{-1}, entonces $x = A^{-1}Ax = A^{-1}b$.

Nota 4 (Importante) ***Suponga que hay un vector x diferente de cero tal que $Ax = \mathbf{0}$.*** *Entonces A no puede tener inversa.* Para repetir: Ninguna matriz es capaz de regresar el 0 a x.

Si A es invertible, entonces $Ax = 0$ sólo puede tener la solución $x = 0$.

Nota 5 Una matriz de 2 por 2 es invertible si y sólo si $ad - bc$ es diferente de cero:

$$\textbf{Inversa de 2 por 2} \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}. \qquad (3)$$

El número $ad - bc$ es el *determinante* de A. Una matriz es invertible si su determinante es distinto de cero (véase el capítulo 4). En MATLAB, la prueba de invertibilidad consiste en *encontrar n pivotes diferentes de cero*. La eliminación produce estos pivotes antes de que aparezca el determinante.

Nota 6 Una matriz diagonal tiene una inversa en el supuesto de que ninguno de los elementos diagonales es cero:

$$\text{Si } A = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d \end{bmatrix} \text{ entonces } A^{-1} = \begin{bmatrix} 1/d_1 & & \\ & \ddots & \\ & & 1/d_n \end{bmatrix} \quad \text{y} \quad AA^{-1} = I.$$

Cuando están implicadas dos matrices, no hay mucho que hacer sobre la inversa de $A + B$. La suma puede o no ser invertible. En vez de lo anterior, la fórmula clave en cálculos matriciales es la inversa de su *producto*, AB. Los números normales son los mismos: $(a + b)^{-1}$ es difícil de simplificar, mientras $1/ab$ se separa en $1/a$ veces $1/b$. Sin embargo, para matrices *el orden de la multiplicación debe ser correcto*: si $ABx = y$, entonces $Bx = A^{-1}y$ y $x = B^{-1}A^{-1}y$. **Las inversas se presentan en orden opuesto.**

1L Un producto AB de matrices invertibles es invertido por $B^{-1}A^{-1}$:

$$\textbf{Inversa de } \boldsymbol{AB} \qquad (AB)^{-1} = B^{-1}A^{-1}. \qquad (4)$$

Demostración Para demostrar que $B^{-1}A^{-1}$ es la inversa de AB, las matrices se multiplican y se aplica la ley asociativa para quitar los paréntesis. Observe que B se escribe al lado de B^{-1}:

$$(AB)(B^{-1}A^{-1}) = ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I$$
$$(B^{-1}A^{-1})(AB) = B^{-1}A^{-1}AB = B^{-1}IB = B^{-1}B = I. \qquad ■$$

Una regla semejante se cumple para tres o más matrices:

$$\textbf{Inversa de } \boldsymbol{ABC} \qquad (ABC)^{-1} = C^{-1}B^{-1}A^{-1}.$$

Este cambio de orden se vio cuando las matrices de eliminación E, F, y G se invirtieron para regresar de U a A. En la dirección hacia adelante, $GFEA$ era U. En la dirección hacia atrás, $L = E^{-1}F^{-1}G^{-1}$ era el producto de las inversas. *Debido a que al último aparece G, entonces G^{-1} aparece primero*. Por favor compruebe que A^{-1} sería $U^{-1}GFE$.

Cálculo de A^{-1}: El método de Gauss-Jordan

Considere la ecuación $AA^{-1} = I$. Si se toma ***una columna a la vez***, la ecuación determina cada columna de A^{-1}. La primera columna de A^{-1} se multiplica por A, para producir la primera columna de la identidad: $Ax_1 = e_1$. De manera semejante, $Ax_2 = e_2$ y $Ax_3 = e_3$; las *es* son las columnas de I. En un ejemplo de 3 por 3, A multiplicada por A^{-1} es I:

$$\boldsymbol{Ax_i = e_i} \qquad \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad (5)$$

Así, se tienen tres sistemas de ecuaciones (o n sistemas). Todos tienen la misma matriz de coeficientes A. Los miembros derechos e_1, e_2, e_3, son diferentes, aunque la eliminación es posible *en todos los sistemas de manera simultánea*. Éste es el ***método de Gauss-Jordan***. En vez de detenerse en U y cambiar a sustitución hacia atrás, continúa restando múltiplos de un renglón *de los renglones de arriba*. Esto produce ceros arriba de la diagonal, y también abajo de ésta. Cuando llega a la matriz identidad, se ha encontrado A^{-1}.

El siguiente ejemplo preserva todas las tres columnas e_1, e_2, e_3, y opera sobre renglones de longitud igual a seis:

Ejemplo 1 Aplicar el método de Gauss-Jordan para encontrar A^{-1}

$$\begin{bmatrix} A & e_1 & e_2 & e_3 \end{bmatrix} = \begin{bmatrix} \mathbf{2} & \mathbf{1} & \mathbf{1} & 1 & 0 & 0 \\ \mathbf{4} & \mathbf{-6} & \mathbf{0} & 0 & 1 & 0 \\ \mathbf{-2} & \mathbf{7} & \mathbf{2} & 0 & 0 & 1 \end{bmatrix}$$

$$\textbf{Pivote} = \mathbf{2} \rightarrow \begin{bmatrix} \mathbf{2} & \mathbf{1} & \mathbf{1} & 1 & 0 & 0 \\ \mathbf{0} & \mathbf{-8} & \mathbf{-2} & -2 & 1 & 0 \\ \mathbf{0} & \mathbf{8} & \mathbf{3} & 1 & 0 & 1 \end{bmatrix}$$

$$\textbf{Pivote} = \mathbf{-8} \rightarrow \begin{bmatrix} \mathbf{2} & \mathbf{1} & \mathbf{1} & 1 & 0 & 0 \\ \mathbf{0} & \mathbf{-8} & \mathbf{-2} & -2 & 1 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & -1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} U & L^{-1} \end{bmatrix}.$$

Esto completa la eliminación hacia adelante, de la primera mitad. La triangular superior U aparece en las tres primeras columnas. Las otras tres columnas son las mismas que en L^{-1}. (Éste es el efecto de aplicar las operaciones elementales GFE a la matriz identidad). Luego, la segunda mitad regresa de U a I (al multiplicar por U^{-1}). Esto lleva L^{-1} a $U^{-1}L^{-1}$, que es A^{-1}. **Creando ceros *arriba* de los pivotes se llega a A^{-1}:**

$$\textbf{Segunda mitad} \begin{bmatrix} U & L^{-1} \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1 & \mathbf{0} & 2 & -1 & -1 \\ 0 & -8 & \mathbf{0} & -4 & 3 & 2 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{bmatrix}$$

$$\textbf{ceros arriba de los pivotes} \rightarrow \begin{bmatrix} 2 & \mathbf{0} & \mathbf{0} & \frac{12}{8} & -\frac{5}{8} & -\frac{6}{8} \\ 0 & -8 & \mathbf{0} & -4 & 3 & 2 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{bmatrix}$$

$$\textbf{se divide entre los pivotes} \rightarrow \begin{bmatrix} 1 & 0 & 0 & \frac{12}{16} & -\frac{5}{16} & -\frac{6}{16} \\ 0 & 1 & 0 & \frac{4}{8} & -\frac{3}{8} & -\frac{2}{8} \\ 0 & 0 & 1 & \mathbf{-1} & \mathbf{1} & \mathbf{1} \end{bmatrix} = \begin{bmatrix} I & A^{-1} \end{bmatrix}.$$

En el último paso, los renglones se dividieron entre sus pivotes 2 y -8 y 1. La matriz de coeficientes en el miembro izquierdo se transformó en la identidad. Debido a que A se convirtió en I, las mismas operaciones en el miembro derecho deben llevar I a A^{-1}. En consecuencia, se ha calculado la inversa.

Una nota para el futuro: El lector puede ver el determinante -16 que aparece en los denominadores de A^{-1}. **El determinante es el producto de los pivotes (2)(-8)(1).** Entra al final cuando los renglones se dividen entre los pivotes.

Observación 1 A pesar de este brillante éxito al calcular A^{-1}, no lo recomiendo. Admito que A^{-1} resuelve en un paso $Ax = b$. Dos pasos triangulares son mejores:

$$x = A^{-1}b \text{ se separa en } Lc = b \text{ y } Ux = c.$$

Puede escribirse $c = L^{-1}b$ y luego $x = U^{-1}c = U^{-1}L^{-1}b$. Sin embargo, observe que estas matrices L^{-1} y U^{-1} no se formaron de manera explícita, y en el cálculo verdadero, *no deben formarse*. Hacerlo sería una pérdida de tiempo, ya que sólo se requiere sustitución hacia atrás para x (y la sustitución hacia adelante produjo c).

Una observación semejante es válida para A^{-1}; la multiplicación $A^{-1}b$ seguiría requiriendo n^2 pasos. ***Lo que se busca es la solución, no todos los elementos de la inversa.***

Observación 2 Sólo por curiosidad, podría contarse el número de operaciones necesarias para encontrar A^{-1}. La cuenta normal para cada nuevo miembro derecho es n^2, la mitad en la dirección hacia adelante y la mitad en la sustitución hacia atrás. Con n miembros derechos $e_1, \ldots, e_n$ esto hace n^3. Luego de incluir las $n^3/3$ operaciones sobre A misma, parece que el total es $4n^3/3$.

Este resultado es ligeramente elevado debido a los ceros en los e_j. La eliminación hacia adelante cambia sólo los ceros que están debajo del 1. Esta parte sólo tiene $n - j$ componentes, de modo que la cuenta para las e_j cambia efectivamente a $(n - j)^2/2$. Al sumar sobre todas las j, el total para la eliminación hacia adelante es $n^3/6$. Esto debe combinarse con las $n^3/3$ operaciones de costumbre que se aplican a A, y los $n(n^2/2)$ pasos de la sustitución hacia atrás que finalmente producen las columnas x_j de A^{-1}. *La cuenta final de multiplicaciones para calcular A^{-1} es n^3:*

$$\textbf{Conteo de operaciones} \qquad \frac{n^3}{6} + \frac{n^3}{3} + n\left(\frac{n^2}{2}\right) = n^3.$$

Este conteo es extraordinariamente bajo. Debido a que la multiplicación de matrices consume n^3 pasos, ¡requiere tantas operaciones para calcular A^2 como las necesarias para calcular A^{-1}! Este hecho parece casi increíble (y calcular A^3 requiere el doble, hasta donde puede verse). A pesar de lo anterior, si A^{-1} no es necesaria, no debe calcularse.

Observación 3 En el cálculo de Gauss-Jordan, siempre se fue en busca de U, antes de iniciar el procedimiento hacia atrás para obtener ceros arriba de los pivotes. Esto es como la eliminación gaussiana, aunque también son posibles otros órdenes. Hubiera podido utilizarse el segundo pivote cuando se estaba ahí antes, con la finalidad de obtener un cero arriba del pivote, así como también abajo de él. Esto no es inteligente. En ese instante el segundo renglón está virtualmente lleno, mientras cerca del extremo tiene ceros provenientes de las operaciones en renglones hacia arriba que ya se habían realizado.

Invertible = No singular (n pivotes)

En última instancia, lo que se quiere saber es cuáles matrices son invertibles y cuáles no lo son. Esta cuestión es tan importante que tiene muchas respuestas. ¡*Consulte la última página del libro*!

En cada uno de los cinco primeros capítulos se proporciona una prueba diferente (aunque equivalente) para comprobar la invertibilidad. Algunas veces las pruebas se extienden a matrices rectangulares e inversas por un lado: el capítulo 2 investiga el tema de renglones y columnas independientes. En el capítulo 3 se invierte AA^T o A^TA. En los otros capítulos se abordan los ***determinantes diferentes de cero***, los ***valores característicos diferentes de cero*** o los ***pivotes diferentes de cero***. Esta última prueba es la que se encuentra en la eliminación gaussiana. Se quiere demostrar (en unos cuantos párrafos teóricos) que la prueba de los pivotes es exitosa.

Suponga que A cuenta con un conjunto completo de pivotes. $AA^{-1} = I$ proporciona n sistemas $Ax_i = e_i$ por separado para las columnas de A^{-1}. Estos sistemas pueden resolverse por eliminación o por Gauss-Jordan. Pueden requerirse intercambios de renglones, pero las columnas de A^{-1} están determinadas.

Hablando estrictamente, es necesario demostrar que la matriz A^{-1} con esas columnas también es *auto*inversa. Al resolver $AA^{-1} = I$ se resuelve al mismo tiempo $A^{-1}A = I$, pero, ¿por qué? **Una inversa por un lado de una matriz cuadrada, es automáticamente una inversa por dos lados.** Para ver por qué, observe que *todo paso de Gauss-Jordan es una multiplicación por la izquierda por una matriz elemental*. Se están permitiendo tres tipos de matrices elementales:

1. E_{ij} para restar un múltiplo ℓ del renglón j del renglón i.
2. P_{ij} para intercambiar los renglones i y j.
3. D (o D^{-1}) para dividir todos los renglones entre sus pivotes.

El proceso de Gauss-Jordan en realidad es una secuencia gigantesca de multiplicaciones de matrices:

$$(D^{-1} \cdots E \cdots P \cdots E)A = I. \tag{6}$$

La matriz entre paréntesis, a la izquierda de A, ¡evidentemente es una autoinversa! Por la nota 2, existe, y es igual a la inversa derecha, de modo que ***toda matriz no singular es invertible***.

La conversa también es verdadera: ***Si A es invertible, tiene n pivotes***. En un caso extremo que es claro: A no puede tener toda una columna de ceros. La inversa jamás podría multiplicar una columna de ceros para producir una columna de I. En un caso menos extremo, suponga que la eliminación inicia en una matriz invertible A, pero que falla en la columna 3:

$$\begin{array}{l}\textbf{Falla}\\ \textbf{En la columna 3 no hay pivote}\end{array} \qquad A' = \begin{bmatrix} d_1 & x & x & x \\ 0 & d_2 & x & x \\ 0 & 0 & 0 & x \\ 0 & 0 & 0 & x \end{bmatrix}.$$

Esta matriz no puede tener una inversa, sin importar cuáles sean las x. Una demostración de este hecho es aplicar operaciones en las columnas (¿por primera vez?) con la finalidad de hacer cero toda la columna. Al restar múltiplos de la columna 2 y luego de la columna 1, se llega a una matriz que ciertamente no es invertible. En consecuencia, la matriz original A no es invertible. La eliminación proporciona una prueba completa: *Una matriz de n por n es invertible si y sólo si tiene n pivotes.*

La matriz traspuesta

Se requiere una matriz más, y por fortuna es mucho más sencilla que la inversa. La **traspuesta** de A se denota por A^T. Sus columnas se toman directamente de los renglones de A: *el i-ésimo renglón de A se convierte en la i-ésima columna de A^T*:

$$\textbf{Traspuesta} \qquad \text{Si } A = \begin{bmatrix} 2 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix} \text{ entonces } A^T = \begin{bmatrix} 2 & 0 \\ 1 & 0 \\ 4 & 3 \end{bmatrix}.$$

Al mismo tiempo, las columnas de A se convierten en los renglones de A^T. Si A es una matriz de m por n, entonces A^T es de n por m. El efecto final es doblar la matriz respecto a su

diagonal principal, y el elemento en el renglón i, columna j de A^T proviene del elemento que está en el renglón j, columna i de A:

$$\textbf{Elementos de } A^T \qquad (A^T)_{ij} = A_{ji}. \qquad (7)$$

La traspuesta de una matriz triangular inferior es triangular superior. La traspuesta de A^T lleva de regreso a A.

Si se suman dos matrices y luego se traspone, el resultado es el mismo que si primero se traspone y luego se suma: $(A + B)^T$ es lo mismo que $A^T + B^T$. Pero, ¿a qué es igual la traspuesta de un producto AB o de una inversa A^{-1}? A continuación se presentan las fórmulas esenciales de esta sección:

1M *i)* La traspuesta de AB es $(AB)^T = B^TA^T$
ii) La traspuesta de A^{-1} es $(A^{-1})^T = (A^T)^{-1}$.

Observe el parecido de la fórmula para $(AB)^T$ con el resultado de la fórmula para $(AB)^{-1}$. En ambos casos se invierte el orden, con lo que se obtiene B^TA^T y $B^{-1}A^{-1}$. La demostración para la inversa fue fácil, pero ésta requiere de una paciencia extraordinaria con la multiplicación de matrices. El primer renglón de $(AB)^T$ es la primera columna de AB. Así, las columnas de A están ponderadas por la primera columna de B. Esto se suma a los renglones de A^T ponderados por el primer renglón de B^T. Eso es exactamente el primer renglón de B^TA^T. Los otros renglones de $(AB)^T$ y B^TA^T también coinciden.

$$\textbf{Se empieza con} \qquad AB = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 3 & 3 \\ 2 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 3 \\ 5 & 5 & 5 \end{bmatrix}$$

$$\textbf{Traspuesta de} \qquad B^TA^T = \begin{bmatrix} 3 & 2 \\ 3 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ 3 & 5 \\ 3 & 5 \end{bmatrix}.$$

Con la finalidad de establecer la fórmula para $(A^{-1})^T$, se inicia con $AA^{-1} = I$ y $A^{-1}A = I$, y se toman las traspuestas. En un miembro, $I^T = I$. En el otro miembro, por el inciso *i)* se conoce la traspuesta de un producto. Puede verse *cómo* $(A^{-1})^T$ es la inversa de A^T, lo cual demuestra *ii)*:

$$\textbf{Inversa de } A^T = \textbf{Traspuesta de } A^{-1} \qquad (A^{-1})^TA^T = I. \qquad (8)$$

Matrices simétricas

Una vez que se han establecido estas reglas, es posible presentar una clase especial de matrices; quizá la clase más importante de todas. ***Una matriz simétrica es una matriz que es igual a su propia traspuesta: $A^T = A$.*** La matriz es necesariamente cuadrada. Cada elemento en un miembro de la diagonal es igual a su "imagen especular" en el otro lado: $a_{ij} = a_{ji}$. Dos ejemplos sencillos son A y D (y también A^{-1}):

$$\textbf{Matrices simétricas} \quad A = \begin{bmatrix} 1 & 2 \\ 2 & 8 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \quad \text{y} \quad A^{-1} = \frac{1}{4}\begin{bmatrix} 8 & -2 \\ -2 & 1 \end{bmatrix}.$$

Una matriz simétrica no necesariamente es invertible; incluso, puede ser una matriz de ceros. *Pero si A^{-1} existe, también es simétrica.* Con base en la fórmula *ii)* anterior, la traspuesta de A^{-1} siempre es igual a $(A^T)^{-1}$; para una matriz simétrica lo anterior es justo A^{-1}. A^{-1} es igual a su propia traspuesta; es simétrica siempre que A lo sea. Ahora se ha demostrado que ***el resultado de multiplicar cualquier matriz R por R^T es una matriz simétrica.***

Productos simétricos R^TR, RR^T, y LDL^T

Sea cualquier matriz R, quizá rectangular. Multiplique R^T por R. Entonces el producto R^TR es automáticamente una matriz simétrica cuadrada:

$$\textit{La traspuesta de } R^TR \textit{ es } R^T(R^T)^T\textbf{, que es } R^TR. \qquad (9)$$

Esta es una demostración rápida de simetría para R^TR. Su elemento i, j es el producto interno del renglón i de R^T (columna i de R) con la columna j de R. El elemento (j, i) es el mismo producto interno, la columna j con la columna i. Así, R^TR es simétrica.

RR^T también es simétrica, aunque diferente de R^TR. En mi experiencia, la mayor parte de los problemas científicos que empiezan con una matriz rectangular R terminan con R^TR, con RR^T o con ambas.

Ejemplo 2 $R = \begin{bmatrix} 1 & 2 \end{bmatrix}$ y $R^T = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ producen $R^TR = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ y $RR^T = \begin{bmatrix} 5 \end{bmatrix}$.

El producto R^TR es de n por n. En orden opuesto, RR^T es de m por m. Incluso si $m = n$, no es muy probable que $R^TR = RR^T$. La igualdad puede ocurrir, aunque no es normal.

Las matrices simétricas se presentan en todo tema cuyas leyes son justas. "Cada acción tiene una reacción igual y opuesta." El elemento a_{ij} que proporciona la acción de i sobre j es compensado por a_{ji}. Esta simetría se verá en la siguiente sección, para ecuaciones diferenciales. Aquí LU pierde la simetría, pero LDL^T la captura perfectamente.

1N Suponga que $A = A^T$ puede factorizarse en $A = LDU$ sin intercambios de renglones. Entonces U es la traspuesta de L. ***La factorización simétrica se vuelve $A = LDL^T$.***

La traspuesta de $A = LDU$ proporciona $A^T = U^TD^TL^T$. Debido a que $A = A^T$, ahora se tienen dos factorizaciones de A en triangular inferior multiplicada por triangular superior (L^T es triangular superior con 1s en la diagonal, exactamente como U.) Debido a que la factorización es única (consulte el problema 17), L^T debe ser idéntica a U.

$$L^T = U \quad \text{y} \quad A = LDL^T \qquad \begin{bmatrix} 1 & 2 \\ 2 & 8 \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{2} & \mathbf{1} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{2} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} = LDL^T.$$

Cuando se aplica eliminación a una matriz simétrica, $A^T = A$ es una ventaja. Las matrices más pequeñas siguen siendo simétricas a medida que procede la eliminación, de modo que ¡Es posible trabajar con la mitad de la matriz! El ángulo inferior del miembro derecho sigue siendo simétrico:

$$\begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \to \begin{bmatrix} a & b & c \\ 0 & d - \dfrac{b^2}{a} & e - \dfrac{bc}{a} \\ 0 & e - \dfrac{bc}{a} & f - \dfrac{c^2}{a} \end{bmatrix}.$$

El trabajo de la eliminación se reduce de $n^3/3$ a $n^3/6$. No es necesario almacenar los elementos de ambos lados de la diagonal, o de almacenar tanto a L como a U.

Conjunto de problemas 1.6

1. Encuentre las inversas (no se requiere ningún sistema especial) de

$$A_1 = \begin{bmatrix} 0 & 2 \\ 3 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 2 & 0 \\ 4 & 2 \end{bmatrix}, \quad A_3 = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}.$$

2. *a)* Encuentre las inversas de las matrices de permutación

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{y} \quad P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

b) Explique para permutaciones, por qué P^{-1} *siempre es igual a* P^{T}. Demuestre que los 1s están en los sitios correctos para obtener $PP^{\mathrm{T}} = I$.

3. A partir de $AB = C$, encuentre una fórmula para A^{-1}. También encuentre A^{-1} a partir de $PA = LU$.

4. *a)* Si A es invertible y $AB = AC$, demuestre rápidamente que $B = C$.

b) Si $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, encuentre un ejemplo con $AB = AC$, pero $B \neq C$.

5. Si la inversa de A^2 es B, demuestre que la inversa de A es AB. (Por tanto, A es invertible siempre que A^2 es invertible.)

6. Aplique el método de Gauss-Jordan para invertir las siguientes matrices

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

7. Encuentre tres matrices de 2 por 2, que no sean $A = I$ y $A = -I$, que sean sus propias inversas: $A^2 = I$.

8. Demuestre que $A = \begin{bmatrix} 1 & 1 \\ 3 & 3 \end{bmatrix}$ no tiene inversa, resolviendo $Ax = 0$, y fallando al resolver

$$\begin{bmatrix} 1 & 1 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

9. Suponga que la eliminación fracasa porque en la columna 3 no hay pivote:

$$\textbf{Pivote faltante} \qquad A = \begin{bmatrix} 2 & 1 & 4 & 6 \\ 0 & 3 & 8 & 5 \\ 0 & 0 & 0 & 7 \\ 0 & 0 & 0 & 9 \end{bmatrix}.$$

Demuestre que A no puede ser invertible. El tercer renglón de A^{-1}, multiplicado por A, debe proporcionar el tercer renglón $[0\ \ 0\ \ 1\ \ 0]$ de $A^{-1}A = I$. ¿Por qué es imposible esto?

10. Encuentre las inversas (de cualquier manera permitida) de

$$A_1 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 3 & 0 & 0 \\ 4 & 0 & 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 \\ 0 & -\frac{2}{3} & 1 & 0 \\ 0 & 0 & -\frac{3}{4} & 1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{bmatrix}.$$

11. Proporcione ejemplos de A y B tales que

a) $A + B$ no es invertible, aunque A y B sí lo son.

b) $A + B$ es invertible, aunque A y B no lo son.

c) Las tres A, B, y $A + B$ son invertibles.

En el último caso su $A^{-1}(A+B)B^{-1} = B^{-1} + A^{-1}$ para demostrar que $C = B^{-1} + A^{-1}$ también es invertible y encuentre una fórmula para C^{-1}.

12. Si A es invertible, ¿qué propiedades de A siguen siendo verdaderas para A^{-1}?

a) A es triangular. *b)* A es simétrica. *c)* A es tridiagonal.

d) Todos los elementos de A son números enteros.

e) Todos los elementos de A son fracciones (incluso los números como $\frac{3}{1}$).

13. Si $A = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ y $B = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, calcule $A^{\mathrm{T}}B$, $B^{\mathrm{T}}A$, AB^{T} y BA^{T}.

14. Si B es cuadrada, demuestre que $A = B + B^{\mathrm{T}}$ siempre es simétrica, y $K = B - B^{\mathrm{T}}$ siempre es *simétrica sesgada*, lo cual significa que $K^{\mathrm{T}} = -K$. Encuentre estas matrices A y K cuando $B = \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix}$, y escriba B como la suma de una matriz simétrica y una matriz simétrica sesgada.

15. *a)* ¿Cuántos elementos pueden elegirse independientemente en una matriz simétrica de orden n?

b) ¿Cuántos elementos pueden elegirse independientemente en una matriz simétrica sesgada ($K^{\mathrm{T}} = -K$) de orden n? ¡La diagonal de K es cero!

16. *a)* Si $A = LDU$, con 1s en la diagonal de L y U, ¿cuál es la factorización correspondiente de A^{T}? Observe que A y A^{T} (matrices cuadradas sin intercambios de renglones) comparten los mismos pivotes.

b) ¿Con qué sistemas triangulares se obtiene la solución de $A^{\mathrm{T}}y = b$?

17. Si $A = L_1D_1U_1$ y $A = L_2D_2U_2$, demuestre que $L_1 = L_2$, $D_1 = D_2$ y $U_1 = U_2$. Si A es invertible, la factorización es única.

a) Deduzca la ecuación $L_1^{-1}L_2D_2 = D_1U_1U_2^{-1}$, y explique por qué un miembro es triangular inferior y el otro es triangular superior.

b) Compare las diagonales principales, y luego compare el resto de las matrices.

18. ¿Cuáles son las condiciones sobre los elementos de A y B para que éstas sean invertibles?

$$A = \begin{bmatrix} a & b & c \\ d & e & 0 \\ f & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & e \end{bmatrix}.$$

19. Compruebe la factorización simétrica LDL^{T} de

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 3 & 12 & 18 \\ 5 & 18 & 30 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}.$$

20. Encuentre la inversa de

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{4} & 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}.$$

21. (Extraordinario) Si A y B son matrices cuadradas, demuestre que $I - BA$ es invertible si $I - AB$ es invertible. Empiece con $B(I - AB) = (I - BA)B$.

22. Encuentre las inversas (directamente o a partir de la fórmula de 2 por 2) de A, B, C:

$$A = \begin{bmatrix} 0 & 3 \\ 4 & 6 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} a & b \\ b & 0 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 3 & 4 \\ 5 & 7 \end{bmatrix}.$$

23. Resuelva para las columnas de $A^{-1} = \begin{bmatrix} x & t \\ y & z \end{bmatrix}$:

$$\begin{bmatrix} 10 & 20 \\ 20 & 50 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 10 & 20 \\ 20 & 50 \end{bmatrix} \begin{bmatrix} t \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

24. Demuestre que $\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ no tiene inversa tratando de resolver para la columna (x, y):

$$\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} x & t \\ y & z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{debe incluir} \quad \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

25. (Importante) Si A es tal que renglón 1 + renglón 2 = renglón 3, demuestre que A no es invertible:
a) Explique por qué $Ax = (1, 0, 0)$ no puede tener una solución.
b) ¿Qué miembros derechos (b_1, b_2, b_3) pueden permitir una solución a $Ax = b$?
c) ¿Qué ocurre al renglón 3 en la eliminación?

26. Si A es tal que columna 1 + columna 2 = columna 3, demuestre que A no es invertible:
a) Encuentre una solución x diferente de cero de $Ax = 0$. La matriz es de 3 por 3.
b) La eliminación preserva columna 1 + columna 2 = columna 3. Explique por qué no hay un tercer pivote.

27. Suponga que A es invertible y que sus dos primeros renglones se intercambian para obtener B. La nueva matriz B, ¿es invertible? ¿Cómo puede obtenerse B^{-1} a partir de A^{-1}?

28. Si el producto $M = ABC$ de tres matrices cuadradas es invertible, entonces A, B, C son invertibles. Encuentre una fórmula para B^{-1} que implique a M^{-1}, A y C.

29. Demuestre que una matriz con una columna de ceros no puede tener una inversa.

30. Multiplique $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ por $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. ¿Cuál es la inversa de cada matriz si $ad \neq bc$?

31. *a*) ¿Qué matriz E tiene el mismo efecto que los siguientes tres pasos? Reste el renglón 1 del renglón 2, reste el renglón 1 del renglón 3, y luego reste el renglón 2 del renglón 3.
b) ¿Qué simple matriz L tiene el mismo efecto que los siguientes tres pasos inversos? Sume el renglón 2 al renglón 3, sume el renglón 1 al renglón 3, y luego sume el renglón 1 al renglón 2.

32. Encuentre los números a y b con los que se obtiene la inversa de 5*eye(4) ones(4, 4):

$$\begin{bmatrix} 4 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 \\ -1 & -1 & 4 & -1 \\ -1 & -1 & -1 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} a & b & b & b \\ b & a & b & b \\ b & b & a & b \\ b & b & b & a \end{bmatrix}.$$

¿Cuáles son a y b en la inversa de 6*eye(5) ones(5, 5)?

33. Demuestre que A = 4*eye(4) – ones(4, 4) *no* es invertible: Multiplique A*ones(4,1).

34. Hay 16 matrices de 2 por 2 cuyos elementos son 1s y 0s. ¿Cuántas de estas matrices son invertibles?

Los problemas 35 a 39 son sobre el método de Gauss-Jordan para calcular A^{-1}.

35. Transforme I en A^{-1} a medida que reduce A a I (por medio de operaciones en los renglones):

$$\begin{bmatrix} A & I \end{bmatrix} = \begin{bmatrix} 1 & 3 & 1 & 0 \\ 2 & 7 & 0 & 1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} A & I \end{bmatrix} = \begin{bmatrix} 1 & 4 & 1 & 0 \\ 3 & 9 & 0 & 1 \end{bmatrix}.$$

36. Imite el ejemplo de 3 por 3 del texto, pero con un signo + en A. Elimine por arriba y por abajo de los pivotes para reducir $[A \ \ I]$ a $[I \ \ A^{-1}]$:

$$\begin{bmatrix} A & I \end{bmatrix} = \begin{bmatrix} \mathbf{2} & \mathbf{1} & \mathbf{0} & 1 & 0 & 0 \\ \mathbf{1} & \mathbf{2} & \mathbf{1} & 0 & 1 & 0 \\ \mathbf{0} & \mathbf{1} & \mathbf{2} & 0 & 0 & 1 \end{bmatrix}.$$

37. Aplique eliminación de Gauss-Jordan sobre $[A \ \ I]$ para resolver $AA^{-1} = I$:

$$\begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} & & \\ x_1 & x_2 & x_3 \\ & & \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

38. Invierta las siguientes matrices por eliminación de Gauss-Jordan, empezando con $[A \ \ I]$:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

39. Intercambie renglones, y prosiga con eliminación de Gauss-Jordan para encontrar A^{-1}:

$$\begin{bmatrix} A & I \end{bmatrix} = \begin{bmatrix} 0 & 2 & 1 & 0 \\ 2 & 2 & 0 & 1 \end{bmatrix}.$$

40. ¿Falso o verdadero? (proporcione un contraejemplo si es falso y una explicación si es verdadero):
a) Una matriz de 4 por 4 con un renglón de ceros no es invertible.
b) Una matriz con 1s abajo de la diagonal principal es invertible.
c) Si A es invertible, entonces A^{-1} es invertible.
d) Si A^{T} es invertible, entonces A es invertible.

41. ¿Para cuáles tres números c la siguiente matriz, no es invertible? ¿Por qué?

$$A = \begin{bmatrix} 2 & c & c \\ c & c & c \\ 8 & 7 & c \end{bmatrix}.$$

42. Demuestre que A es invertible si $a \neq 0$ y $a \neq b$ (encuentre los pivotes y A^{-1}):

$$A = \begin{bmatrix} a & b & b \\ a & a & b \\ a & a & a \end{bmatrix}.$$

43. La inversa de la siguiente matriz es extraordinaria. Encuéntrela por eliminación sobre $[A\ \ I]$. Extienda lo anterior a una "matriz alternante" de 5 por 5, y conjeture su inversa:

$$A = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

44. Si B tiene las columnas de A en orden invertido, resuelva $(A - B)x = 0$ para demostrar que $A - B$ no es invertible. Un ejemplo lo conducirá a x.

45. Encuentre y compruebe las inversas (suponiendo que existen) de las siguientes matrices en bloque:

$$\begin{bmatrix} I & 0 \\ C & I \end{bmatrix} \quad \begin{bmatrix} A & 0 \\ C & D \end{bmatrix} \quad \begin{bmatrix} 0 & I \\ I & D \end{bmatrix}.$$

46. Use inv(S) para invertir la matriz simétrica MATLAB de 4 por 4 S = pascal(4). Obtenga la matriz Pascal triangular inferior A = abs(pascal(4,1) y test inv(S) = inv(A')* inv(A).

47. Si A = ones(4,4) y b = rand(4,1), ¿cómo le hace saber MATLAB que $Ax = b$ no tiene solución? Si b = ones(4,1), ¿qué solución de $Ax = b$ se encuentra por medio de $A\backslash b$?

48. M^{-1} muestra el cambio en A^{-1} (lo cual es útil saber) cuando una matriz se resta de A. Compruebe el siguiente inciso 3 cuidadosamente, multiplicando MM^{-1} para obtener I:

1. $M = I - uv^T$ y $M^{-1} = I + uv^T/(1 - v^Tu)$.
2. $M = A - uv^T$ y $M^{-1} = A^{-1} + A^{-1}uv^TA^{-1}/(1 - v^TA^{-1}u)$.
3. $M = I - UV$ y $M^{-1} = I_n + U(I_m - VU)^{-1}V$.
4. $M = A - UW^{-1}V$ y $M^{-1} = A^{-1} + A^{-1}U(W - VA^{-1}U)^{-1}VA^{-1}$.

Las cuatro identidades provienen del bloque 1, 1 luego de que se invierten las siguientes matrices:

$$\begin{bmatrix} I & u \\ v^T & 1 \end{bmatrix} \quad \begin{bmatrix} A & u \\ v^T & 1 \end{bmatrix} \quad \begin{bmatrix} I_n & U \\ V & I_m \end{bmatrix} \quad \begin{bmatrix} A & U \\ V & W \end{bmatrix}.$$

Los problemas 49 a 55 son sobre las reglas para trasponer matrices.

49. Encuentre A^T y A^{-1} y $(A^{-1})^T$ y $(A^T)^{-1}$ para

$$A = \begin{bmatrix} 1 & 0 \\ 9 & 3 \end{bmatrix} \text{ y también } A = \begin{bmatrix} 1 & c \\ c & 0 \end{bmatrix}.$$

50. Compruebe que $(AB)^T$ es igual a B^TA^T, aunque éstas son diferentes de A^TB^T:

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \quad AB = \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}.$$

En caso de que $AB = BA$ (¡lo cual en general no es cierto!), ¿cómo puede demostrar que $B^TA^T = A^TB^T$?

51. *a)* La matriz $((AB)^{-1})^T$ proviene de $(A^{-1})^T$ y $(B^{-1})^T$. *¿En qué orden?*

b) Si U es triangular superior, entonces $(U^{-1})^T$ es _____ triangular.

52. Demuestre que $A^2 = 0$ es posible pero $A^TA = 0$ no es posible (a menos que A = matriz cero).

53. *a)* Al multiplicar el vector renglón x^T por la columna y, ¿qué número se obtiene?

$$x^TAy = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = ___.$$

b) Este es el renglón x^TA = _____ veces la columna $y = (0, 1, 0)$.

c) Este es el renglón $x^T = [0\ \ 1]$ veces la columna Ay = _____.

54. Cuando se traspone una matriz en bloque $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, el resultado es M^T = _____. Pruébelo. ¿Qué condiciones deben imponerse a A, B, C, D para que la matriz en bloque sea simétrica?

55. Explique por qué el producto interno de x y y es igual al producto interno de Px y Py. Luego, $(Px)^T(Py) = x^Ty$, indica que $P^TP = I$ para cualquier permutación. Con $x = (1, 2, 3)$ y $y = (1, 4, 2)$, escoja P para demostrar que $(Px)^Ty$ no es siempre igual a $x^T(P^Ty)$.

Los problemas 56 a 60 son sobre matrices simétricas, y sus factorizaciones.

56. Si $A = A^T$ y $B = B^T$, ¿cuáles de las siguientes matrices son ciertamente simétricas?

a) $A^2 - B^2$ *b)* $(A + B)(A - B)$ *c)* ABA *d)* $ABAB$

57. Si $A = A^T$ requiere un intercambio de renglones, entonces también necesita un intercambio de columnas para permanecer simétrica. En lenguaje de matrices, PA pierde la simetría de A pero _____ recupera la simetría.

58. *a)* ¿Cuántos elementos de A pueden elegirse independientemente, si $A = A^T$ es de 5 por 5?

b) ¿De qué manera pueden L y D (de 5 por 5) proporcionar el mismo número de opciones en LDL^T?

59. Suponga que R es rectangular (de m por n) y que A es simétrica (de m por m).

a) Trasponga R^TAR para mostrar su simetría. ¿De qué forma es esta matriz?

b) Demuestre por qué R^TR no contiene números negativos en su diagonal.

60. Factorice las siguientes matrices en $A = LDL^T$. La matriz D es diagonal:

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 1 & b \\ b & c \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}.$$

Los tres siguientes problemas son sobre aplicaciones de $(Ax)^Ty = x^T(A^Ty)$.

61. Las ciudades de Boston, Chicago y Seattle están conectadas con cables que conducen energía eléctrica. Los voltajes que reciben estas ciudades son x_B, x_C, x_S. Con resistencias unitarias entre las ciudades, las tres corrientes están en y:

$$y = Ax \quad \text{es} \quad \begin{bmatrix} y_{BC} \\ y_{CS} \\ y_{BS} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_B \\ x_C \\ x_S \end{bmatrix}.$$

a) Encuentre las corrientes totales A^Ty de las tres ciudades.

b) Compruebe que $(Ax)^Ty$ coincide con $x^T(A^Ty)$: hay seis términos en cada una.

62. La producción de x_1 camiones y x_2 aviones requiere $x_1 + 50x_2$ toneladas de acero, $40x_1 + 1000x_2$ libras de caucho, y $2x_1 + 50x_2$ meses de trabajo. Si los costos unita-

rios y_1, y_2, y_3 son \$700 por tonelada, \$3 por libra, y \$3000 por mes, ¿cuáles son los valores de un camión y un avión? Éstas son las componentes de $A^{\mathrm{T}}y$.

63. Ax proporciona las cantidades de acero, caucho y trabajo para producir x en el problema 62. Encuentre A. Luego $(Ax)^{\mathrm{T}}y$ es el ______ de entradas, mientras $x^{\mathrm{T}}(A^{\mathrm{T}}y)$ es el valor de ______.

64. A continuación se presenta otra factorización de A en una *triangular multiplicada por una simétrica:*

Se empieza con $A = LDU$. Luego, A es igual a $L(U^{\mathrm{T}})^{-1}$ por $U^{\mathrm{T}}DU$.

¿Por qué es triangular $L(U^{\mathrm{T}})^{-1}$? En su diagonal sólo hay 1s. ¿Por qué es simétrica $U^{\mathrm{T}}DU$?

65. Un *grupo* de matrices incluye a AB y A^{-1} si incluye a A y B. "Los productos y las inversas permanecen en grupos." ¿Cuáles de los siguientes conjuntos son grupos? Las matrices triangulares inferiores L con 1s en su diagonal, las matrices simétricas S, las matrices positivas M, las matrices diagonales invertibles D, las matrices de permutación P. Invente dos grupos de matrices más.

66. Si todo renglón de una matriz de 4 por 4 contiene a los números 0, 1, 2, 3 en algún orden, ¿la matriz puede ser simétrica? ¿Puede ser invertible?

67. Demuestre que ningún reordenamiento de renglones ni de columnas puede trasponer una matriz típica.

68. Una ***matriz noroeste*** cuadrada B es cero en la esquina sureste, abajo de la antidiagonal que une $(1, n)$ con $(n, 1)$. B^{T} y B^2, ¿son matrices noroeste? B^{-1}, ¿es noroeste o sureste? ¿Cuál es la forma de BC = ***noroeste multiplicada por sureste***? Se permite combinar permutaciones con las L y U de costumbre (suroeste y noreste).

69. Compare tic; inv(A); toc para A = rand(500) y A = rand(1000). El conteo n^3 indica que el tiempo de cómputo (medido por tic; toc) debe multiplicarse por 8 cuando n se duplica. ¿Cree el lector que esta A aleatoria es invertible?

70. I = eye(1000); A = rand(1000); B = triu(A); produce una matriz *triangular* aleatoria B. Compare los tiempos para inv(B) y $B\backslash I$. La diagonal al revés se somete a ingeniería con la finalidad de utilizar los ceros en B, mientras inv utiliza los ceros en I cuando $[B\ \ I]$ se reduce por Gauss-Jordan. (También compare con inv(A) y $A\backslash I$ para toda la matriz A.)

71. Demuestre que L^{-1} tiene elementos j/i para $i \leq j$ (la matriz $-1, 2, -1$ tiene esta L):

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 \\ 0 & -\frac{2}{3} & 1 & 0 \\ 0 & 0 & -\frac{3}{4} & 1 \end{bmatrix} \quad \text{y} \quad L^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & 1 & 0 \\ \frac{1}{4} & \frac{2}{4} & \frac{3}{4} & 1 \end{bmatrix}.$$

Pruebe este patrón para L = eye(5) − diag(1:5)\diag(1:4,−1) e inv(L).

1.7 MATRICES ESPECIALES Y APLICACIONES

Esta sección tiene dos objetivos. El primero consiste en explicar una forma en la que grandes sistemas lineales $Ax = b$ pueden presentarse en la práctica. La verdad es que un problema grande y realista en ingeniería o economía nos llevaría demasiado lejos del alcance de este libro, aunque hay una aplicación natural e importante que no requiere mucha preparación.

El otro objetivo consiste en ilustrar, aprovechando la misma aplicación, las propiedades especiales que suelen tener las matrices de coeficientes. Las matrices grandes siempre presentan un patrón definido; a menudo un patrón de simetría, y *muchísimos elementos iguales a cero*. Debido a que una matriz "holgada" contiene mucho menos de n^2 piezas de información, los cálculos deben ser rápidos. Se abordarán las *matrices banda*, con la finalidad de ver cómo la concentración cerca de la diagonal acelera la eliminación. De hecho, se analiza una matriz tridiagonal especial.

La matriz misma puede verse en la ecuación (6). Proviene del intercambio de una ecuación diferencial con una ecuación matricial. El problema continuo pide por $u(x)$ en cada x, por lo que una computadora no es capaz de resolverlo exactamente. Debe ser aproximado por un problema discreto: mientras más incógnitas se tienen, mejor es la precisión y mayor el gasto. Como un problema continuo sencillo aunque también bastante típico, nuestra elección se dirige a la ecuación diferencial

$$-\frac{d^2u}{dx^2} = f(x), \quad 0 \leq x \leq 1. \tag{1}$$

Esta es una ecuación lineal para la función incógnita $u(x)$. Cualquier combinación $C + Dx$ puede sumarse a cualquier solución, ya que la segunda derivada de $C + Dx$ no contribuye con nada. La incertidumbre que dejan estas dos constantes arbitrarias C y D se elimina mediante una "*condición a la frontera*" en cada extremo del intervalo:

$$u(0) = 0, \qquad u(1) = 0. \tag{2}$$

El resultado es un *problema de dos puntos con valor en la frontera*, que describe un fenómeno que no es transitorio, sino de estado estacionario; por ejemplo, la distribución de temperatura en una varilla cuyos extremos están fijos a 0°, y con una fuente de calor $f(x)$.

Recuerde que nuestro objetivo es producir un problema discreto; en otras palabras, un problema en álgebra lineal. Es por ello que sólo es posible aceptar una cantidad finita de información sobre $f(x)$, por ejemplo, sus valores en n puntos equidistantes $x = h, x = 2h, \ldots, x = nh$. Para la solución verdadera u se calculan valores aproximados $u_1, \ldots, u_n$ en esos mismos puntos. En los extremos $x = 0$ y $x = 1 = (n + 1)h$, los valores en la frontera son $u_0 = 0$ y $u_{n+1} = 0$.

La primera pregunta es: ¿Cómo se sustituye la derivada d^2u/dx^2? La primera derivada puede ser aproximada deteniendo $\Delta u/\Delta x$ en un tamaño de paso finito, y no permitiendo que h (o Δx) tienda a cero. La diferencia Δu puede ser *hacia adelante*, *hacia atrás*, o *centrada*:

$$\frac{\mathbf{\Delta u}}{\mathbf{\Delta x}} = \frac{u(x+h) - u(x)}{h} \ \text{o bien,}\ \frac{u(x) - u(x-h)}{h} \ \text{o bien,}\ \frac{u(x+h) - u(x-h)}{2h}. \tag{3}$$

La última expresión es simétrica respecto a x, y es la más precisa. Para la segunda derivada hay justo una combinación que sólo utiliza los valores de x y $x \pm h$:

Segunda diferencia
$$\frac{d^2u}{dx^2} \approx \frac{\mathbf{\Delta^2 u}}{\mathbf{\Delta x^2}} = \frac{u(x+h) - 2u(x) + u(x-h)}{h^2}. \tag{4}$$

Lo anterior también tiene el mérito de ser simétrico con respecto a x. Para repetir, el miembro derecho tiende al valor verdadero de d^2u/dx^2 cuando $h \to 0$, aunque es necesario detenerse en una h positiva.

En cada punto de la malla, $x = jh$, la ecuación $-d^2u/dx^2 = f(x)$ se sustituye por su análogo discreto (5). Se multiplicó por h^2 para alcanzar n ecuaciones $Au = b$:

Ecuación en diferencias
$$-u_{j+1} + 2u_j - u_{j-1} = h^2 f(jh) \quad \text{para} \quad j = 1, \ldots, n. \tag{5}$$

Las ecuaciones primera y última ($j = 1$ y $j = n$) incluyen $u_0 = 0$ y $u_{n+1} = 0$, que se conocen a partir de las condiciones a la frontera. Estos valores deben desplazarse al miembro derecho de la ecuación si se desea que sean distintos de cero. La estructura de estas n ecua-

ciones (5) puede visualizarse mejor en forma matricial. Se escoge $h = \frac{1}{6}$ para obtener una matriz A de 5 por 5:

Ecuación matricial
$$\begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{bmatrix} = h^2 \begin{bmatrix} f(h) \\ f(2h) \\ f(3h) \\ f(4h) \\ f(5h) \end{bmatrix}. \qquad (6)$$

A partir de este momento *se trabajará con la ecuación* (6). Tiene una matriz de coeficientes bastante regular, cuyo orden n puede ser muy grande. La matriz A tiene muchas propiedades especiales, de las cuales tres son fundamentales:

1. ***La matriz A es tridiagonal.*** Todos los elementos diferentes de cero están en la diagonal principal y en las dos diagonales adyacentes. Fuera de esta banda todos los elementos son $a_{ij} = 0$. Estos ceros traerán una simplificación considerable para la eliminación gaussiana.
2. ***La matriz es simétrica.*** Cada elemento a_{ij} es igual a su imagen especular a_{ji}, de modo que $A^T = A$. La triangular superior U es la traspuesta de la triangular inferior L, y $A = LDL^T$. Esta simetría de A refleja la simetría de d^2u/dx^2. Una derivada impar como du/dx o d^3u/dx^3 destruiría la simetría.
3. ***La matriz es positiva definida.*** Esta propiedad adicional indica que *los pivotes son positivos*. Los intercambios de renglones son innecesarios en la teoría y en la práctica. Esto contrasta con la matriz B, al final de esta sección, que no es positiva definida. Sin ningún intercambio de renglones, es totalmente vulnerable al redondeo.
   La característica de positiva definida reúne todo este curso (¡en el capítulo 6!)

Ahora se vuelve al hecho de que A es tridiagonal. ¿Cuál es el efecto de esto en la eliminación? La primera etapa del proceso de eliminación produce ceros abajo del primer pivote:

Eliminación en A: Paso 1
$$\begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} \to \begin{bmatrix} 2 & -1 & & & \\ 0 & \frac{3}{2} & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix}.$$

En comparación con una matriz general de 5 por 5, este paso presenta dos simplificaciones importantes:

1. Abajo del pivote *sólo hay un elemento diferente de cero.*
2. El renglón del pivote es *muy corto.*

El multiplicador $\ell_{21} = -\frac{1}{2}$ proviene de una división. El nuevo pivote, $\frac{3}{2}$, proviene de una *simple multiplicación-sustracción*. Además, *se preserva el patrón tridiagonal*: Toda etapa de la eliminación acepta las simplificaciones *a*) y *b*).

El resultado final es la factorización $LDU = LDL^T$ de A. ¡Observe los pivotes!

$$A = \begin{bmatrix} 1 & & & & \\ -\frac{1}{2} & 1 & & & \\ & -\frac{2}{3} & 1 & & \\ & & -\frac{3}{4} & 1 & \\ & & & -\frac{4}{5} & 1 \end{bmatrix} \begin{bmatrix} \frac{2}{1} & & & & \\ & \frac{3}{2} & & & \\ & & \frac{4}{3} & & \\ & & & \frac{5}{4} & \\ & & & & \frac{6}{5} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & & & \\ & 1 & -\frac{2}{3} & & \\ & & 1 & -\frac{3}{4} & \\ & & & 1 & -\frac{4}{5} \\ & & & & 1 \end{bmatrix}.$$

Los factores L y U de una matriz tridiagonal son bidiagonales. Los tres factores juntos tienen la misma estructura de banda de tres diagonales esenciales ($3n - 2$ parámetros) que A. También observe que L y U son traspuestas entre sí, como era de esperar por la simetría. Todos los pivotes 2/1, 3/2, 4/3, 5/4, 6/5 son positivos. Su producto es el ***determinante*** de A: $\det A = 6$. Resulta evidente que los pivotes convergen a 1 cuando n se hace grande. Estas matrices hacen muy felices a las computadoras.

Estos factores holgados L y U cambian por completo el conteo de operaciones de costumbre. La eliminación en cada columna requiere sólo dos operaciones, como arriba, y hay n columnas. *En vez de las $n^3/3$ operaciones, sólo se necesitan* $2n$. Los sistemas tridiagonales $Ax = b$ pueden resolverse casi de inmediato. *El costo de resolución de un sistema tridiagonal es proporcional a n.*

Una **matriz de banda** tiene $a_{ij} = 0$ excepto en la banda $|i - j| < w$ (véase la figura 1.8). El "semiancho de banda" es $w = 1$ para una matriz diagonal, $w = 2$ para una matriz tridiagonal, y $w = n$ para una matriz completa. Para cada columna, la eliminación requiere $w(w - 1)$ operaciones: un renglón de longitud w actúa sobre $w - 1$ abajo. *La eliminación en las n columnas de una matriz de banda requiere alrededor de w^2n operaciones.*

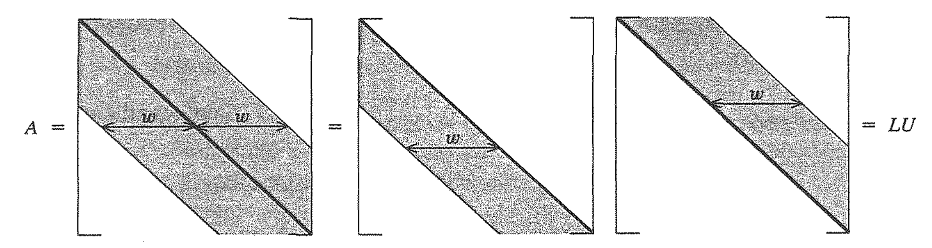


Figura 1.8 Una matriz de banda A y sus factores L y U.

Cuando w tiende a n, la matriz se vuelve completa, y el conteo es aproximadamente n^3. Para un conteo exacto, la esquina inferior del miembro derecho ya no tiene espacio para un ancho de banda w. El número preciso de divisiones y multiplicaciones-sustracciones que producen L, D, U (sin suponer que A es simétrica) es $P = \frac{1}{3}w(w-1)(3n - 2w + 1)$. Para una matriz completa con $w = n$, se recupera $P = \frac{1}{3}n(n-1)(n+1)$. Este es un número entero, ya que $n - 1$, n, y $n + 1$ son enteros consecutivos, y uno de ellos es divisible entre 3.

Este es el último conteo de operaciones, y se recalca la cuestión más importante. Una matriz de diferencias finitas como A tiene una inversa completa. Al resolver $Ax = b$, *en realidad se está en una peor situación si se conoce A^{-1} que si se conocen* L y U. La multiplicación de A^{-1} por b requiere n^2 pasos, mientras para la eliminación hacia adelante y la sustitución hacia atrás que producen $x = U^{-1}c = U^{-1}L^{-1}b = A^{-1}b$, son suficientes $4n$ pasos.

Esperamos que este ejemplo haya reforzado la comprensión que el lector tiene sobre la eliminación (¡de la cual ahora suponemos que está perfectamente entendida!) Se trata de un ejemplo genuino de los grandes sistemas lineales que en realidad se encuentran en la práctica. En el siguiente capítulo la atención se centra en la existencia y unicidad de x, para m ecuaciones en n incógnitas.

Error por redondeo

En teoría, el caso no singular se ha completado. Hay un conjunto completo de pivotes (con intercambios de renglones). En la práctica, *más intercambios de renglones* pueden ser igualmente necesarios, o bien, la solución calculada puede convertirse fácilmente en inútil.

Dedicaremos dos páginas (completamente opcionales en clase) para hacer más estable la eliminación: por qué es necesaria y cómo se lleva a cabo.

Para un sistema de tamaño moderado, por ejemplo de 100 por 100, la eliminación implica alrededor de 330 000 operaciones ($\frac{1}{3}n^3$). Con cada operación es necesario esperar un error por redondeo. Normalmente, se mantiene fijo un número de dígitos significativos (por ejemplo tres para una computadora extremadamente débil). Así, la adición de dos números de tamaños diferentes proporciona un error:

Error por redondeo $0.456 + 0.00123 \rightarrow 0.457$ pierde los dígitos 2 y 3.

¿Cómo contribuyen todos estos errores individuales al error final en $Ax = b$?

Este problema no es fácil. Fue atacado por John von Neumann, quien era el matemático más reputado en la época en que las computadoras repentinamente hacían posible un millón de operaciones. De hecho, la combinación de Gauss y Von Neumann proporciona al simple algoritmo de eliminación una historia extraordinariamente distinguida, aunque incluso Von Neumann sobreestimó el error por redondeo final. Quien encontró la forma correcta para contestar la pregunta fue Wilkinson, cuyos libros se han convertido en clásicos.

Dos ejemplos sencillos ilustrarán tres cuestiones importantes sobre el error por redondeo. Los ejemplo son

Mal acondicionado $A = \begin{bmatrix} 1. & 1. \\ 1. & 1.0001 \end{bmatrix}$ **Bien acondicionado** $B = \begin{bmatrix} 0.0001 & 1. \\ 1. & 1. \end{bmatrix}$.

A es casi singular, mientras B está lejos de ser singular. Si el último elemento de A se modifica ligeramente a $a_{22} = 1$, entonces A *es* singular. Considere dos miembros derechos bastante parecidos:

$$\begin{aligned} u + \quad\; v &= 2 \\ u + 1.0001v &= 2 \end{aligned} \quad \text{y} \quad \begin{aligned} u + \quad\; v &= 2 \\ u + 1.0001v &= 2.0001 \end{aligned}$$

La solución del primero es $u = 2$, $v = 0$. La solución del segundo es $u = v = 1$. *Un cambio en el quinto dígito de b fue amplificado a un cambio en el primer dígito de la solución. Ningún método numérico es capaz de evitar esta sensibilidad a pequeñas perturbaciones.* El mal acondicionamiento puede desplazarse de un lado a otro, pero no es posible eliminarlo. La verdadera solución es muy sensible, y la solución calculada no puede dejar de serlo.

El segundo punto es como sigue.

1O Incluso una matriz bien acondicionada como B puede ser arruinada por un algoritmo deficiente.

Lamentamos decir que para la matriz B, la eliminación gaussiana directa es un algoritmo deficiente. Suponga que .0001 es aceptado como el primer pivote. Luego, 10 000 veces el primer renglón se resta del segundo. El elemento inferior derecho se convierte en –9999, pero el redondeo hasta tres cifras proporcionaría –10 000. Cualquier traza del elemento 1 desaparecería:

Eliminación en B con un pivote pequeño $$\begin{aligned} 0.0001u + v &= 1 \\ u + v &= 2 \end{aligned} \longrightarrow \begin{aligned} 0.0001u + v &= 1 \\ -9999v &= -9998. \end{aligned}$$

El redondeo produce $-10\,000v = -10\,000$, o bien $v = 1$. Esto es correcto hasta tres cifras decimales. La sustitución hacia atrás con el $v = .9999$ correcto proporciona $u = 1$:

Resultado correcto $0.0001u + 0.9999 = 1$, o bien, $u = 1$.

En vez de lo anterior, al aceptar $v = 1$, que es erróneo sólo en la cuarta cifra decimal, se obtiene $u = 0$:

Resultado erróneo $0.0001u + 1 = 1$, o bien, $u = 0$.

La u calculada es completamente errónea. B está bien acondicionada pero la eliminación es violentamente inestable. L, D, y U están completamente fuera de escala con B:

$$B = \begin{bmatrix} 1 & 0 \\ 10\,000 & 1 \end{bmatrix} \begin{bmatrix} 0.0001 & 0 \\ 0 & -9999 \end{bmatrix} \begin{bmatrix} 1 & 10\,000 \\ 0 & 1 \end{bmatrix}.$$

El pequeño pivote 0.0001 trajo insensibilidad, y el remedio es evidente: *intercambiar renglones*.

1P Un pivote pequeño obliga a un cambio práctico en la eliminación. Normalmente, cada pivote se compara con todos los posibles pivotes en la misma columna. Intercambiar renglones para obtener el pivote más grande posible se denomina ***pivoteo parcial***.

Para B, el pivote 0.0001 podría compararse con el pivote posible 1 que está abajo de él. Un intercambio de renglones tendría lugar de inmediato. En términos matriciales, esta es una multiplicación por una matriz de permutación $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. La nueva matriz $C = PB$ tiene buenos factores:

$$C = \begin{bmatrix} 1 & 1 \\ 0.0001 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0.0001 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0.9999 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Los pivotes para C son 1 y 0.9999, mucho mejores que 0.0001 y -9999 para B.

La estrategia del *pivoteo completo* también busca en todas las últimas columnas el pivote más grande posible. No sólo podría ser necesario un intercambio de renglones, sino también uno de columnas. (Esto es la *pos*multiplicación, por una matriz de permutación). La dificultad al ser tan conservador es el gasto, y el pivoteo parcial es bastante adecuado.

Finalmente se ha llegado al algoritmo fundamental del álgebra lineal numérica: la ***eliminación con pivoteo parcial***. Algunos refinamientos adicionales, como, ver si es necesario reescalar todo un renglón o toda una columna siguen siendo posibles. Pero esencialmente, el lector sabe ahora lo que hace una computadora con un sistema de ecuaciones lineales. En comparación con esta descripción "teórica" *—encontrar A^{-1}, y multiplicar $A^{-1}b$*— nuestra descripción ha consumido bastante del tiempo (y paciencia) del lector. Desearía que hubiese una forma mas fácil de explicar cómo x se encuentra realmente, aunque no creo que la haya.

Conjunto de problemas 1.7

1. Escriba los factores $LDU = LDL^T$ de A en la ecuación (6) cuando $n = 4$. Encuentre el determinante como el producto de los pivotes en D.

2. Modifique a_{11} en la ecuación (6) de $a_{11} = 2$ a $a_{11} = 1$, y encuentre los factores LDU de esta nueva matriz tridiagonal.

3. Encuentre la matriz A_0 de 5 por 5 ($h = \frac{1}{6}$) que aproxima

$$-\frac{d^2u}{dx^2} = f(x), \qquad \frac{du}{dx}(0) = \frac{du}{dx}(1) = 0,$$

sustituyendo estas condiciones a la frontera por $u_0 = u_1$ y $u_6 = u_5$. Compruebe que su A_0 multiplicada por el vector constante (C, C, C, C, C) produce cero; A_0 *es singular*. De manera semejante, si $u(x)$ es una solución del problema continuo, entonces también lo es $u(x) + C$.

4. Escriba la ecuación de la matriz de diferencias finitas de 3 por 3 ($h = \frac{1}{4}$) para

$$-\frac{d^2u}{dx^2} + u = x, \qquad u(0) = u(1) = 0.$$

5. Con $h = \frac{1}{4}$ y $f(x) = 4\pi^2 \operatorname{sen} 2\pi x$, la ecuación en diferencias (5) es

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \frac{\pi^2}{4} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

Resuelva para u_1, u_2, u_3 y encuentre su error en comparación con la solución verdadera $u = \operatorname{sen} 2\pi x$ en $x = \frac{1}{4}$, $x = \frac{1}{2}$, y $x = \frac{3}{4}$.

6. ¿Cuál es el sistema de 5 por 5 que sustituye a (6) si las condiciones en la frontera se cambian a $u(0) = 1$, $u(1) = 0$?

Los problemas 7 a 11 son sobre el error por redondeo y los intercambios de renglones.

7. Calcule H^{-1} en dos formas para la matriz de Hilbert de 3 por 3

$$H = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix},$$

primero por cálculo exacto y luego redondeando cada número hasta tres cifras. Esta matriz H está mal acondicionada y los intercambios de renglones no son de utilidad.

8. Para la misma matriz H, compare los miembros derechos de $Hx = b$ cuando las soluciones son $x = (1, 1, 1)$ y $x = (0, 6, -3.6)$.

9. Resuelva $Hx = b = (1, 0, \ldots, 0)$ para la matriz de Hilbert de 10 por 10 con $h_{ij} = 1/(i + j - 1)$, usando cualquier código de computadora para ecuaciones lineales. Luego, intercambie un elemento de b por .0001, y compare las soluciones.

10. Compare los pivotes en eliminación directa con los del pivoteo parcial para la siguiente matriz.

$$A = \begin{bmatrix} 0.001 & 0 \\ 1 & 1000 \end{bmatrix}$$

(En realidad, éste es un ejemplo que requiere reescalamiento antes de la eliminación.)

11. Explique por qué el pivoteo parcial produce multiplicadores ℓ_{ij} en L que satisfacen $|\ell_{ij}| \le 1$. ¿Puede proporcionar un ejemplo de 3 por 3 con todos los $|a_{ij}| \le 1$ cuyo último pivote sea 4? Este es el peor de los casos, ya que cada elemento es cuando mucho multiplicado por 2 cuando $|\ell_{ij}| \le 1$.

Capítulo 1 Ejercicios de repaso

1.1 *a*) Escriba las matrices de 3 por 3 con elementos

$$a_{ij} = i - j \quad \text{y} \quad b_{ij} = \frac{i}{j}.$$

b) Calcule los productos AB, BA y A^2.

1.2 Para las matrices

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix},$$

calcule AB y BA y A^{-1} y B^{-1} y $(AB)^{-1}$.

1.3 Encuentre ejemplos de matrices de 2 por 2 con $a_{12} = \frac{1}{2}$ para las cuales

a) $A^2 = I$. *b*) $A^{-1} = A^{\mathrm{T}}$. *c*) $A^2 = A$.

1.4 Resuelva por eliminación y sustitución hacia atrás.

$$\begin{aligned} u \quad\;\; + w &= 4 \\ u + v \quad\;\; &= 3 \\ u + v + w &= 6 \end{aligned} \qquad \text{y} \qquad \begin{aligned} v + w &= 0 \\ u \quad\;\; + w &= 0 \\ u + v \quad\;\; &= 6. \end{aligned}$$

1.5 Factorice las matrices precedentes en $A = LU$ o bien $PA = LU$.

1.6 *a*) Hay 16 matrices de 2 por 2 cuyos elementos son 1s y 0s. ¿Cuántas son invertibles?

b) (¡Mucho más difícil!) Si al azar se escriben 1s y 0s en los elementos de una matriz de 10 por 10, ¿qué es más probable: que la matriz sea invertible o que sea singular?

1.7 Hay 16 matrices de 2 por 2 cuyos elementos son 1s y –1s. ¿Cuántas son invertibles?

1.8 ¿Cómo están relacionados los renglones de EA con los renglones de A en los casos siguientes?

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 4 & 0 & 1 \end{bmatrix} \text{ o bien, } E = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \text{ o bien, } E = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

1.9 Escriba un sistema de 2 por 2 que tenga una infinidad de soluciones.

1.10 Encuentre inversas, si existen, por inspección o por eliminación de Gauss-Jordan:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 1 & 1 & -2 \\ 1 & -2 & 1 \\ -2 & 1 & 1 \end{bmatrix}.$$

1.11 Si E es de 2 por 2 y suma la primera ecuación a la segunda, ¿cuáles son las matrices E^2 y E^8 y $8E$?

1.12 ¿Falso o verdadero? Proporcione una *explicación* si es verdadero y un *contraejemplo* si es falso:

1) Si A es invertible y sus renglones aparecen en orden invertido en B, entonces B es invertible.

2) Si A y B son simétricas, entonces AB es simétrica.
3) Si A y B son invertibles, entonces BA es invertible.
4) Toda matriz no singular puede factorizarse en el producto $A = LU$ de una triangular inferior L y una triangular superior U.

1.13 Resuelva $Ax = b$ resolviendo los sistemas triangulares $Lc = b$ y $Ux = c$:

$$A = LU = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

¿Qué parte de A^{-1} encontró con esa b particular?

1.14 De ser posible, encuentre matrices B de 3 por 3 tales que

a) $BA = 2A$ para toda A.
b) $BA = 2B$ para toda A.
c) BA tenga los renglones primero y último de A, invertidos.
d) BA tenga las columnas primera y última de A, invertidas.

1.15 Encuentre el valor de c en la siguiente inversa de n por n.

$$\text{si } A = \begin{bmatrix} n & -1 & \cdot & -1 \\ -1 & n & \cdot & -1 \\ \cdot & \cdot & \cdot & -1 \\ -1 & -1 & -1 & n \end{bmatrix} \text{ entonces } A^{-1} = \frac{1}{n+1} \begin{bmatrix} c & 1 & \cdot & 1 \\ 1 & c & \cdot & 1 \\ \cdot & \cdot & \cdot & 1 \\ 1 & 1 & 1 & c \end{bmatrix}.$$

1.16 ¿Para qué valores de k el sistema

$$\begin{aligned} kx + y &= 1 \\ x + ky &= 1 \end{aligned}$$

no tiene solución, tiene una, solución, o bien tiene una infinidad de soluciones?

1.17 Encuentre la factorización simétrica $A = LDL^T$ de

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 6 & 4 \\ 0 & 4 & 11 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

1.18 Suponga que A es la matriz identidad de 4 por 4, excepto por un vector v en la columna 2:

$$A = \begin{bmatrix} 1 & v_1 & 0 & 0 \\ 0 & v_2 & 0 & 0 \\ 0 & v_3 & 1 & 0 \\ 0 & v_4 & 0 & 1 \end{bmatrix}.$$

a) Factorice A en LU, suponiendo $v_2 \neq 0$.
b) Encuentre A^{-1}, que tiene la misma forma que A.

1.19 Resuelva por eliminación, o demuestre que no hay solución:

$$\begin{aligned} u + v + w &= 0 \\ u + 2v + 3w &= 0 \\ 3u + 5v + 7w &= 1 \end{aligned} \quad \text{y} \quad \begin{aligned} u + v + w &= 0 \\ u + v + 3w &= 0 \\ 3u + 5v + 7w &= 1. \end{aligned}$$

1.20 Las matrices de permutación de n por n constituyen un ejemplo importante de un "grupo". Si se multiplican, se sigue permaneciendo dentro del grupo: tienen inversas en el grupo; la identidad está en el grupo; y la ley $P_1(P_2P_3) = (P_1P_2)P_3$ es verdadera, ya que es verdadera para todas las matrices.

a) ¿Cuántos miembros pertenecen a los grupos de las matrices de permutación de 4 por 4 y de n por n?
b) Encuentre una potencia k de modo que todas las matrices de permutación de 3 por 3 cumplan $P^k = I$.

1.21 Describa los renglones de DA y las columnas de AD si $D = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$.

1.22 *a)* Si A es invertible, ¿cuál es la inversa de A^T?
b) Si A también es simétrica, ¿cuál es la traspuesta de A^{-1}?
c) Ilustre ambas fórmulas cuando $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$.

1.23 Experimente con $n = 2$ y $n = 3$, para encontrar

$$\begin{bmatrix} 2 & 3 \\ 0 & 0 \end{bmatrix}^n, \quad \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}^n, \quad \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}^{-1}.$$

1.24 Empiece con un primer plano $u + 2v - w = 6$, para encontrar la ecuación de lo siguiente:

a) El plano paralelo que pasa por el origen.
b) Un segundo plano que también contenga los puntos (6, 0, 0) y (2, 2, 0).
c) Un tercer plano que se corte con los planos primero y segundo en el punto (4, 1, 0).

1.25 ¿Qué múltiplo del renglón 2 se resta del renglón 3 en la eliminación hacia adelante de A?

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix}.$$

¿Cómo sabe el lector (sin multiplicar estos factores) que A es *invertible, simétrica,* y *tridiagonal*? ¿Cuáles son los pivotes?

1.26 *a)* ¿Cuál es el vector x que hace Ax = columna 1 de A + 2(columna 3), para una matriz A de 3 por 3?
b) Obtenga una matriz que cumpla columna 1 + 2(columna 3) = 0. Compruebe que A es singular (menos de 3 pivotes), y explique por qué, éste debe ser el caso.

1.27 ¿Falso o verdadero? Proporcione un contraejemplo si es falso, y una explicación si es verdadero:

1) Si $L_1U_1 = L_2U_2$ (las matrices U son triangulares superiores con diagonal diferente de cero, y las matrices L son triangulares inferiores con diagonal de 1s), entonces $L_1 = L_2$ y $U_1 = U_2$. La factorización LU, es única.
2) Si $A^2 + A = I$, entonces $A^{-1} = A + I$.
3) Si todos los elementos en la diagonal de A son cero, entonces A es singular.

1.28 Al tanteo o con la eliminación de Gauss-Jordan calcule

$$\begin{bmatrix} 1 & 0 & 0 \\ \ell & 1 & 0 \\ m & 0 & 1 \end{bmatrix}^n, \quad \begin{bmatrix} 1 & 0 & 0 \\ \ell & 1 & 0 \\ m & 0 & 1 \end{bmatrix}^{-1}, \quad \begin{bmatrix} 1 & 0 & 0 \\ \ell & 1 & 0 \\ 0 & m & 1 \end{bmatrix}^{-1}.$$

1.29 Escriba las matrices de 2 por 2, tales que

a) Invierten la dirección de cada vector.

b) Proyectan cada vector sobre el eje x_2.

c) Hacen girar 90° a cada vector en sentido contrario al sentido de movimiento de las manecillas del reloj.

d) Reflejan cada vector a 45° respecto a la recta $x_1 = x_2$.

Capítulo



Espacios vectoriales

2.1 ESPACIOS Y SUBESPACIOS VECTORIALES

La eliminación puede simplificar, elemento por elemento, el sistema lineal $Ax = b$. Por fortuna, también simplifica la teoría. Las preguntas básicas de *existencia* y *unicidad*: ¿existe una solución?, ¿no existe ninguna solución?, o bien, ¿existe una infinidad de soluciones? Todas estas preguntas podrán responderse más fácilmente después de aplicar la eliminación. Para tratar estas interrogantes se requiere una sección más, con la finalidad de encontrar todas las soluciones para un sistema de m por n. Así el ciclo de ideas estará completo.

Sin embargo, la eliminación sólo produce un tipo de entendimiento sobre $Ax = b$. Nuestro objetivo principal es lograr una comprensión distinta y más profunda. Se trata del corazón del álgebra lineal.

Para la noción de ***espacio vectorial***, de inmediato se empieza con los espacios más importantes, que se denotan por $\mathbf{R}^1$, $\mathbf{R}^2$, $\mathbf{R}^3$, . . . ; el espacio $\mathbf{R}^n$ *consta de todos los vectores columna con n componentes*. (Se escribe $\mathbf{R}$ porque los componentes son números reales). $\mathbf{R}^2$ se representa por el plano x-y de costumbre; las dos componentes del vector se convierten en las coordenadas x y y del punto correspondiente. Las tres componentes de un vector en $\mathbf{R}^3$ proporcionan un punto en el espacio tridimensional. El espacio unidimensional $\mathbf{R}^1$ es una recta.

Lo importante para el álgebra lineal es que la extensión a n dimensiones es directa. Para un vector en $\mathbf{R}^7$ sólo se requieren siete componentes, incluso si es difícil visualizar la geometría. En todos los espacios vectoriales son posibles dos operaciones:

> ***Es posible sumar dos vectores cualesquiera, y todos los vectores pueden multiplicarse por escalares.***
>
> **En otras palabras, puede trabajarse con combinaciones lineales.**

La suma obedece la ley conmutativa $x + y = y + x$; existe un "vector cero" que cumple $0 + x = x$; y hay un vector "$-x$" que satisface $-x + x = 0$. Ocho propiedades (incluyendo las tres anteriores) son fundamentales: la lista completa se proporciona en el problema 5 al final de esta sección. ***Un espacio vectorial real es un conjunto de vectores junto con reglas para la suma vectorial y la multiplicación por números reales.*** La suma y la multiplicación deben producir vectores en el espacio, y deben cumplir las ocho condiciones.

Normalmente, estos ocho vectores pertenecen a uno de los espacios $\mathbf{R}^n$; son vectores columna normales. Si $x = (1, 0, 0, 3)$, entonces las componentes de $2x$ (y también de $x + x$) son 2, 0, 0, 6. La definición formal permite que otros entes sean "vectores", en el supuesto de que la adición y la multiplicación por escalares estén bien. Se proporcionan tres ejemplos:

1. *El espacio dimensional infinito* $\mathbf{R}^\infty$. Sus vectores tienen una infinidad de componentes, como en $x = (1, 2, 1, 2, \ldots)$. Las leyes para $x + y$ y cx permanecen sin cambio.
2. *El espacio de las matrices de 3 por 2*. ¡En este caso los "vectores" son matrices! Es posible sumar dos matrices, y $A + B = B + A$; existe una matriz cero, etc. Este espacio es casi el mismo que $\mathbf{R}^6$. (Las seis componentes están dispuestas en un rectángulo, en vez de en una columna.) Cualquier elección de m y n proporcionaría, como ejemplo semejante, el espacio vectorial de todas las matrices de m por n.
3. *El espacio de las funciones* $f(x)$. Aquí se admiten todas las funciones f que están definidas en un intervalo fijo, por ejemplo $0 \le x \le 1$. El espacio incluye a $f(x) = x^2$, $g(x) = \operatorname{sen} x$, su suma $(f + g)(x) = x^2 + \operatorname{sen} x$, y todos los múltiplos como $3x^2$ y $-\operatorname{sen} x$. Los vectores son funciones, y la dimensión es de alguna manera un infinito más grande que para $\mathbf{R}^\infty$.

En los ejercicios se proporcionan otros ejemplos, pero los espacios vectoriales que más requerimos, están en otra parte: ***se encuentran dentro de los espacios estándares*** $\mathbf{R}^n$. Deseamos describirlos y explicar por qué son importantes. Geométricamente, piense en el espacio tridimensional de costumbre $\mathbf{R}^3$ y elija cualquier plano que pase por el origen. ***Ese plano es un espacio vectorial por derecho propio.*** Si un vector en el plano se multiplica por 3, o por -3, se obtiene un vector en el mismo plano. Si se suman dos vectores en el plano, su suma permanece en el plano. Este plano que pasa por $(0, 0, 0)$ ilustra uno de los conceptos más importantes en el álgebra lineal; se trata de un ***subespacio*** del espacio original $\mathbf{R}^3$.

DEFINICIÓN Un ***subespacio*** de un espacio vectorial es un conjunto no vacío que satisface los requisitos de un espacio vectorial: ***las combinaciones lineales permanecen en el subespacio.***

i) Si se suman dos vectores cualesquiera en el subespacio, *$x + y$ está en el subespacio.*

ii) Si cualquier vector x en el subespacio se multiplica por cualquier escalar c, *cx está en el subespacio.*

Observe cómo se recalca la palabra ***espacio***. Un sub*espacio* es un subconjunto "cerrado" bajo la suma y la multiplicación por un escalar. Estas operaciones cumplen las reglas del espacio huésped, ya que ***se sigue estando dentro del subespacio.*** Las ocho propiedades requeridas siguen cumpliéndose en el espacio más grande, por lo que se cumplen automáticamente en todo subespacio. Observe en particular que ***el vector cero pertenece a todo subespacio.*** Este hecho se debe a la regla ii): el escalar se escoge como $c = 0$.

El subespacio más pequeño $\mathbf{Z}$ contiene sólo un vector: el vector cero. Se trata de un "espacio con dimensión cero", que sólo contiene el punto en el origen. Las reglas i) y ii) se cumplen, ya que la suma $0 + 0$ está en este espacio con un punto, así como todos los multiplicadores $c0$. *Este es el espacio vectorial más pequeño posible*: el espacio vacío no está permitido. En el otro extremo, el subespacio más grande es todo el espacio original. Si el espacio original es $\mathbf{R}^3$, entonces es fácil describir los subespacios posibles: $\mathbf{R}^3$ mismo, cualquier plano que pase por el origen, cualquier recta que pase por el origen, o sólo el origen (el vector cero).

La diferencia entre un subconjunto y un subespacio se aclarará mediante ejemplos. En cada caso, ¿es posible sumar vectores, y multiplicar por escalares sin salir del espacio?

Ejemplo 1 Considere todos los vectores en $\mathbf{R}^2$ cuyas componentes son positivas o cero. Este subconjunto es el primer cuadrante del plano x-y; las coordenadas satisfacen $x \ge 0$ y $y \ge 0$. *No es un subespacio*, aunque contiene al cero, y la adición permanece dentro del subconjunto. Se viola la regla ii), ya que si el escalar es -1 y el vector es $[1\ \ 1]$, entonces el múltiplo $cx = [-1\ \ -1]$ está en el tercer cuadrante, no en el primero.

Si se incluye el tercer cuadrante junto con el primero, la multiplicación por un escalar está bien. Todo múltiplo cx permanece en este subconjunto. No obstante, ahora se viola la regla i), ya que al sumar $[1\ 2] + [-2\ -1]$ se obtiene $[-1\ 1]$, que no está en ninguno de los cuadrantes mencionados. El menor subespacio que contiene al primer cuadrante es todo el espacio $\mathbf{R}^2$.

Ejemplo 2 Empiece por considerar el espacio vectorial de todas las matrices de 3 por 3. Un subespacio posible es el conjunto de las *matrices triangulares inferiores*. Otro es el conjunto de las *matrices simétricas*. $A + B$ y cA son triangulares inferiores si A y B son triangulares inferiores, y son simétricas si A y B son simétricas. Por supuesto, la matriz cero, está en ambos subespacios.

El espacio columna de *A*

A continuación se proporcionan los ejemplos clave: el **espacio columna** y el **espacio nulo** de una matriz A. ***El espacio columna contiene a todas las combinaciones lineales de las columnas de A.*** Es un subespacio de $\mathbf{R}^m$. Se ilustra con un sistema de $m = 3$ ecuaciones en $n = 2$ incógnitas:

La combinación de las columnas es igual a *b*
$$\begin{bmatrix} 1 & 0 \\ 5 & 4 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \qquad (1)$$

Con $m > n$ se tienen más ecuaciones que incógnitas; en cuyo caso *normalmente no hay solución*. El sistema sólo es resoluble para un subconjunto muy "delgado" de todas las bs posibles. Una forma para describir este subconjunto delgado, es tan simple que es fácil ignorarla.

2A El sistema $Ax = b$ es resoluble si y sólo si el vector b puede expresarse como una combinación de las columnas de A. Por tanto, b está en el espacio columna.

Esta descripción sólo implica un replanteamiento de $Ax = b$ *por columnas*:

Combinación de columnas
$$u \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix} + v \begin{bmatrix} 0 \\ 4 \\ 4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \qquad (2)$$

Estas son las mismas tres ecuaciones en dos incógnitas. Ahora el problema es: encontrar números u y v que multiplican a las columnas primera y segunda para producir b. El sistema es resoluble exactamente cuando estos coeficientes existen, y el vector (u, v) es la solución x.

Se está afirmando que los miembros derechos obtenibles *b son todas las combinaciones de las columnas de A*. Un miembro derecho posible es la columna en sí; los pesos son $u = 1$ y $v = 0$. Otra posibilidad es la segunda columna: $u = 0$ y $v = 1$. Otra tercera posibilidad es el miembro derecho $b = 0$. Con $u = 0$ y $v = 0$, el vector $b = 0$ siempre puede obtenerse.

Es posible describir *todas las combinaciones* de las dos columnas geométricamente; *$Ax = b$ puede resolverse si y sólo si b está en el* ***plano*** *generado por los dos vectores columna* (véase la figura 2.1). Este es el conjunto delgado de bs obtenibles. Si b está fuera del plano, entonces no es una combinación de las dos columnas. En ese caso $Ax = b$ no tiene solución.

Lo importante es que este plano no sólo es un subconjunto de $\mathbf{R}^3$; es un subespacio. Se trata del ***espacio columna*** de A, que consta de ***todas las combinaciones lineales de las columnas***. Se denota por $C(A)$. Es fácil comprobar los requerimientos i) y ii) para un subespacio de $\mathbf{R}^m$:

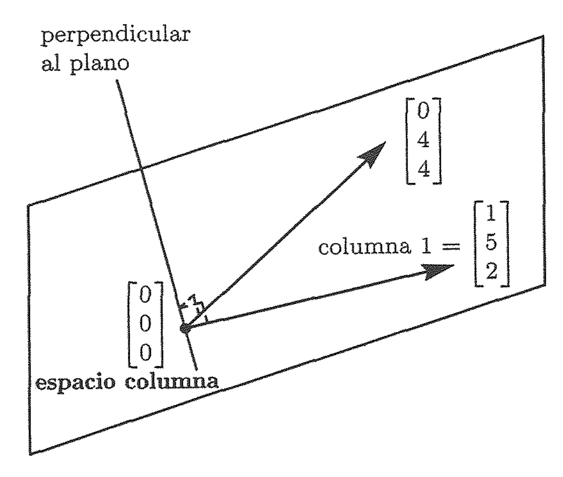


Figura 2.1 El espacio columna $C(A)$, un plano en el espacio tridimensional.

i) Suponga que b y b' están en el espacio columna, de modo que $Ax = b$ para alguna x y $Ax' = b'$ para alguna x'. Luego, $A(x + x') = b + b'$, de modo que $b + b'$ también es una combinación de las columnas. El espacio columna de todos los vectores obtenibles b es cerrado bajo la suma.

ii) Si b está en el espacio columna $C(A)$, también lo está cualquier múltiplo cb. Si alguna combinación de columnas produce b (por ejemplo $Ax = b$), entonces al multiplicar esa combinación por c produce cb. En otras palabras, $A(cx) = cb$.

Para otra matriz A, las dimensiones de la figura 2.1 pueden ser muy distintas. El espacio columna más pequeño posible (con sólo un vector) proviene de la matriz cero $A = 0$. La única combinación de las columnas es $b = 0$. En el otro extremo, suponga que A es la matriz identidad de 5 por 5. Entonces $C(I)$ es todo $\mathbf{R}^5$; las cinco columnas de I pueden combinarse para producir cualquier vector pentadimensional b. Esto no es en absoluto especial de la matriz identidad. *Cualquier matriz de 5 por 5 que sea no singular tiene como espacio columna todo* $\mathbf{R}^5$. Para una matriz así, $Ax = b$ puede resolverse por eliminación gaussiana; hay cinco pivotes. En consecuencia, todo b está en $C(A)$ para una matriz no singular.

Ahora puede darse cuenta de cómo el capítulo 1 está contenido en este capítulo. En aquél se estudiaron matrices de n por n cuyo espacio columna es $\mathbf{R}^n$. Ahora se permiten matrices singulares, y matrices rectangulares de cualquier forma. Así, $C(A)$ puede estar en alguna parte entre el espacio cero y todo el espacio $\mathbf{R}^m$. Junto con este espacio perpendicular, constituye uno de nuestros dos métodos para comprender $Ax = b$.

El espacio nulo de A

El segundo método para tratar $Ax = b$ es "dual" al primero. No sólo se tiene interés en los miembros derechos obtenibles b, sino también en las soluciones x que los obtienen. El miembro derecho $b = 0$ siempre permite la solución $x = 0$, aunque puede haber una infinidad de otras soluciones. (Siempre hay, en caso de que haya más incógnitas que ecuaciones, $n > m$). ***Las soluciones de*** $Ax = 0$ ***constituyen un espacio vectorial: el espacio nulo de A.***

> El ***espacio nulo*** de una matriz consta de todos los vectores x tales que $Ax = 0$. Se denota por $N(A)$. Es un subespacio de $\mathbf{R}^n$, justo como el espacio columna era un subespacio de $\mathbf{R}^m$.

Se cumple el requerimiento i): Si $Ax = 0$ y $Ax' = 0$, entonces $A(x + x') = 0$. El requerimiento ii) también se cumple: Si $Ax = 0$ entonces $A(cx) = 0$. ¡Ambos requerimientos fallan si el miembro derecho no es cero! Sólo las soluciones de una ecuación *homogénea* ($b = 0$) constituyen un subespacio. Es fácil encontrar el espacio nulo para el ejemplo que acaba de proporcionarse; es lo más pequeño posible:

$$\begin{bmatrix}1 & 0\\5 & 4\\2 & 4\end{bmatrix}\begin{bmatrix}u\\v\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$

La primera ecuación proporciona $u = 0$, y entonces la segunda ecuación obliga a que $v = 0$. El espacio nulo sólo contiene al vector (0, 0). Esta matriz tiene "columnas independientes", un concepto clave que se presentará en breve.

La situación cambia cuando una tercera columna es una combinación de las dos primeras:

$$\textbf{Espacio nulo más grande} \qquad B = \begin{bmatrix}1 & 0 & 1\\5 & 4 & 9\\2 & 4 & 6\end{bmatrix}.$$

B tiene el mismo espacio columna que A. La nueva columna está en el plano de la figura 2.1; es la suma de los dos vectores columna con los que se empezó. Sin embargo, el espacio nulo de B contiene al vector $(1, 1, -1)$, por lo que automáticamente contiene a cualquier múltiplo $(c, c, -c)$:

$$\textbf{El espacio nulo es una recta} \qquad \begin{bmatrix}1 & 0 & 1\\5 & 4 & 9\\2 & 4 & 6\end{bmatrix}\begin{bmatrix}c\\c\\-c\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$

El espacio nulo de B es la recta que contiene a todos los puntos $x = c$, $y = c$, $z = -c$. (La recta pasa por el origen, como debe hacer cualquier subespacio). Lo que se desea es poder encontrar, para cualquier sistema $Ax = b$, y encontrar $C(A)$ y $N(A)$: todos los miembros derechos obtenibles b, y todas las soluciones de $Ax = 0$.

Los vectores b, están en el espacio columna y los vectores x están en el espacio nulo. Se calcularán las dimensiones de estos espacios, y un conjunto idóneo de vectores para generarlos. Esperamos terminar comprendiendo todos los *cuatro* subespacios que están estrechamente relacionados entre sí y con A: el espacio columna de A, el espacio nulo de A, y sus dos espacios perpendiculares.

Conjunto de problemas 2.1

1. Construya un subconjunto del plano x-y $\mathbf{R}^2$ que sea
   a) cerrado bajo la suma y resta, pero no bajo la multiplicación por un escalar.
   b) cerrado bajo la multiplicación por un escalar, pero no bajo la suma vectorial.
   Sugerencia: Empiece con u y v; sume y reste para el inciso *a*). Intente cu y cv para el inciso *b*).

2. ¿Cuáles de los siguientes subconjuntos de R^3 son realmente subespacios?
   a) El plano de vectores (b_1, b_2, b_3) cuya primera componente es $b_1 = 0$.

b) El plano de vectores b con $b_1 = 1$.
c) Los vectores b con $b_2 b_3 = 0$ (esta es la unión de dos subespacios, el plano $b_2 = 0$ y el plano $b_3 = 0$).
d) Todas las combinaciones de dos vectores dados (1, 1, 0) y (2, 0, 1).
e) El plano de vectores (b_1, b_2, b_3) que satisface $b_3 - b_2 + 3b_1 = 0$.

3. Describa el espacio columna, y el espacio nulo de las matrices

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0 & 0 & 3 \\ 1 & 2 & 3 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

4. ¿Cuál es el menor subespacio de matrices de 3 por 3 que contiene a todas las matrices simétricas, *y a* todas las matrices triangulares inferiores? ¿Cuál es el mayor subespacio que está contenido en cada uno de estos dos subespacios?

5. Se requiere que la suma y la multiplicación por un escalar cumplan las ocho siguientes condiciones:

1. $x + y = y + x$.
2. $x + (y + z) = (x + y) + z$.
3. Hay un "vector cero" único tal que $x + 0 = x$ para toda x.
4. Para cada x hay un vector único $-x$ tal que $x + (-x) = 0$.
5. $1x = x$.
6. $(c_1 c_2)x = c_1(c_2 x)$.
7. $c(x + y) = cx + cy$.
8. $(c_1 + c_2)x = c_1 x + c_2 x$.

a) Suponga que la adición en $\mathbf{R}^2$ suma un 1 extra a cada componente, de modo que (3, 1) + (5, 0) es igual a (9, 2) en vez de (8, 1). Si la multiplicación por un escalar permanece sin cambio, ¿qué reglas se rompen?
b) Demuestre que el conjunto de todos los números reales positivos con $x + y$ y cx vueltos a definir como de costumbre, igual a xy y x^c, es un espacio vectorial. ¿Cuál es el "vector cero"?
c) Suponga que $(x_1, x_2) + (y_1, y_2)$ se define como $(x_1 + y_2, x_2 + y_1)$. Con $cx = (cx_1, cx_2)$ de costumbre, ¿cuáles de las ocho condiciones no se cumplen?

6. Sea **P** el plano en tres dimensiones con ecuación $x + 2y + z = 6$. ¿Cuál es la ecuación del plano $\mathbf{P}_0$ que pasa por el origen y es paralelo a **P**? ¿**P** y $\mathbf{P}_0$, son subespacios de $\mathbf{R}^3$?

7. ¿Cuáles de los siguientes son subespacios de $\mathbf{R}^\infty$?

a) Todas las sucesiones como (1, 0, 1, 0, . . .) que incluyen una infinidad de ceros.
b) Todas las sucesiones $(x_1, x_2, \ldots)$ con $x_j = 0$ a partir de un punto.
c) Todas las sucesiones decrecientes: $x_{j+1} \leq x_j$ para cada j.
d) Todas las sucesiones convergentes: la x_j tiene límite cuando $j \to \infty$.
e) Todas las progresiones aritméticas: $x_{j+1} - x_j$ es la misma para toda j.
f) Todas las progresiones geométricas $(x_1, kx_1, k^2x_1, \ldots)$ permitiendo toda k y x_1.

8. ¿Cuáles de las siguientes descripciones con correctas? Las soluciones x de

$$Ax = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

constituyen

a) Un plano.
b) Una recta.
c) Un punto.
d) Un subespacio.
e) El espacio nulo de A.
f) El espacio columna de A.

9. Demuestre que el conjunto de matrices no singulares de 2 por 2 no es un espacio vectorial. También demuestre que el conjunto de matrices *singulares* de 2 por 2 no es un espacio vectorial.

10. La matriz $A = \begin{bmatrix} 2 & -2 \\ 2 & -2 \end{bmatrix}$ es un "vector" en el espacio **M** de todas las matrices de 2 por 2. Escriba el vector cero en este espacio, el vector $\frac{1}{2}A$ y el vector $-A$. ¿Qué matrices están en el menor subespacio que contiene a A?

11. *a*) Describa un subespacio de **M** que contenga a $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ pero no a $B = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$.
b) Si un subespacio de **M** contiene a A y a B, ¿debe contener a I?
c) Describa un subespacio de **M** que contenga matrices diagonales diferentes de cero.

12. Las funciones $f(x) = x^2$ y $g(x) = 5x$ son "vectores" en el espacio vectorial **F** de todas las funciones reales. La combinación $3f(x) - 4g(x)$ es la función $h(x) =$ ____. ¿Cuál regla se rompe si al multiplicar $f(x)$ por c se obtiene la función $f(cx)$?

13. Si la suma de los "vectores" $f(x)$ y $g(x)$ en **F** se define como $f(g(x))$, entonces el "vector cero" es $g(x) = x$. Considere la multiplicación por un escalar de costumbre $cf(x)$, y encuentre dos reglas que no se cumplen.

14. Describa el menor subespacio del espacio vectorial **M** de las matrices de 2 por 2 que contiene

a) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ y $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. *b*) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ y $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

c) $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$. *d*) $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$.

15. Sea **P** el plano en $\mathbf{R}^3$ con ecuación $x + y - 2z = 4$. ¡El origen (0, 0, 0) no está en **P**! Encuentre dos vectores en **P**, y compruebe que su suma no está en **P**.

16. $\mathbf{P}_0$ es el plano que pasa por (0, 0, 0) y es paralelo al plano **P** del problema 15. ¿Cuál es la ecuación de $\mathbf{P}_0$? Encuentre dos vectores en $\mathbf{P}_0$, y compruebe que su suma está en $\mathbf{P}_0$.

17. Los cuatro tipos de subespacios de $\mathbf{R}^3$ son planos, rectas, $\mathbf{R}^3$ mismo, o **Z** que sólo contiene a (0, 0, 0).
a) Describa los tres tipos de subespacios de $\mathbf{R}^2$.
b) Describa los cinco tipos de subespacios de $\mathbf{R}^4$.

18. a) La intersección de dos planos que pasan por (0, 0, 0) probablemente es una ____, aunque puede ser un ____. ¡No puede ser el vector cero **Z**!
b) La intersección de un plano que pasa por (0, 0, 0) con una recta que pasa por (0, 0, 0) probablemente es un ____, aunque puede ser un ____.
c) Si **S** y **T** son subespacios de $\mathbf{R}^5$, su intersección $\mathbf{S} \cap \mathbf{T}$ (vectores en ambos subespacios) es un subespacio de $\mathbf{R}^5$. *Compruebe los requerimientos sobre* $x + y$ y cx.

19. Suponga que **P** es un plano que pasa por (0, 0, 0) y que **L** es una recta que pasa por (0, 0, 0). El menor espacio vectorial que contiene tanto a **P** como a **L** es ____ o ____.

20. ¿Falso o verdadero? para **M** = todas las matrices de 3 por 3 (compruebe la suma utilizando un ejemplo).
a) Las matrices simétricas sesgadas en **M** (con $A^T = -A$) constituyen un subespacio.
b) Las matrices simétricas en **M** (con $A^T \neq A$) constituyen un subespacio.
c) Las matrices que tienen a (1, 1, 1) en su espacio nulo constituyen un subespacio.

Los problemas 21 a 30 son sobre espacios columna $C(A)$, y la ecuación $Ax = b$.

21. Describa los espacios columna (rectas o planos) de estas matrices particulares:

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

22. ¿Para qué miembros derechos (encuentre una condición sobre b_1, b_2, b_3) los siguientes sistemas son resolubles?

$$a) = \begin{bmatrix} 1 & 4 & 2 \\ 2 & 8 & 4 \\ -1 & -4 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \qquad b) \begin{bmatrix} 1 & 4 \\ 2 & 9 \\ -1 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

23. Al sumar el renglón 1 de A al renglón 2 se obtiene B. Al sumar la columna 1 a la columna 2 se obtiene C. Una combinación de las columnas de ____ también es una combinación de las columnas de A. De las siguientes matrices, ¿cuáles tienen la misma columna ____?

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}.$$

24. ¿Para qué vectores (b_1, b_2, b_3) los siguientes sistemas tienen una solución?

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

25. (Recomendado) Si a una matriz A se suma una columna extra b, entonces el espacio columna se vuelve más grande, a menos que ____. Proporcione un ejemplo en el que el espacio columna se vuelva más grande, y un ejemplo en que no lo haga. ¿Por qué $Ax = b$ es resoluble exactamente cuando el espacio columna *no se vuelve* más grande al incluir a b?

26. Las columnas de $A\,B$ son una combinación de las columnas de A. Esto significa: *El espacio columna de $A\,B$ está contenido en* (y quizá es igual a) *el espacio columna de A*. Proporcione un ejemplo en el que los espacios columna de A y AB no sean iguales.

27. Si A es cualquier matriz invertible de 8 por 8, entonces su espacio columna es ____. ¿Por qué?

28. ¿Falso o verdadero? (proporcione un contraejemplo si es falso).
a) Los vectores b que no están en el espacio columna $C(A)$ constituyen un subespacio.
b) Si $C(A)$ contiene sólo al vector cero, entonces A es la matriz cero.
c) El espacio columna de $2A$ es igual al espacio columna de A.
d) El espacio columna de $A - I$ es igual al espacio columna de A.

29. Construya una matriz de 3 por 3 cuyo espacio columna contenga a (1, 1, 0) y a (1, 0, 1) pero no a (1, 1, 1). Construya una matriz de 3 por 3 cuyo espacio columna sea sólo una recta.

30. Si el sistema de 9 por 12 $Ax = b$ es resoluble para toda b, entonces $C(A)$ = ____.

31. ¿Por qué $\mathbf{R}^2$ no es un subespacio de $\mathbf{R}^3$?

2.2 CÓMO RESOLVER $Ax = 0$ Y $Ax = b$

El capítulo 1 se centró en matrices invertibles cuadradas. Había una solución de $Ax = b$, que era $x = A^{-1}b$. Esta solución se encontró por eliminación (y no calculando A^{-1}). Una matriz rectangular trae nuevas posibilidades: U puede no tener un conjunto completo de pivotes. Esta sección aborda el tema de ir de U a una forma reducida R: **la matriz más simple que puede obtenerse con eliminación**. R revela de inmediato todas las soluciones.

Para una matriz invertible, el espacio nulo sólo contiene a $x = 0$ (multiplicado $Ax = 0$ por A^{-1}). El espacio columna es todo el espacio ($Ax = b$ no tiene solución para toda b). Las nuevas preguntas surgen cuando el espacio nulo contiene *más que al vector cero* y/o el espacio columna contiene *menos que todos los vectores*:

1. Cualquier vector x_n en el espacio nulo puede sumarse a una solución particular x_p. Las soluciones de todas las ecuaciones lineales tienen esta forma: $x = x_p + x_n$:

   Solución completa $Ax_p = b$, y $Ax = 0$ producen $A(x_p + x_n) = b$.

2. Cuando el espacio columna no contiene a toda b en $\mathbf{R}^m$, se requieren las condiciones sobre b para hacer resoluble a $Ax = b$.

Un ejemplo de 3 por 4 es un tamaño aceptable. Se escribirán todas las soluciones de $Ax = 0$. Se encontrarán las condiciones para que b esté en el espacio columna (de modo que $Ax = b$ sea resoluble). El sistema de 1 por 1 $0x = b$, con una ecuación y una incógnita, presenta dos posibilidades:

$0x = b$ *no tiene soluciones* a menos que $b = 0$. El espacio columna de la matriz cero de 1 por 1 sólo contiene a $b = 0$.

$0x = 0$ *tiene una infinidad de soluciones*. El espacio nulo contiene a *toda* x. Una solución particular es $x_p = 0$, y la solución completa es $x = x_p + x_n = 0 +$ (cualquier x).

Admito que es sencillo. Si se pasa a una matriz de 2 por 2, la cuestión se vuelve más interesante. La matriz $\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$ no es invertible: $y + z = b_1$ y $2y + 2z = b_2$ suele no tener solución.

No hay solución a menos que $b_2 = 2b_1$. El espacio columna de A sólo contiene bs que son múltiplos de (1, 2).

Cuando $b_2 = 2b_1$, ***hay una infinidad de soluciones***. Una solución particular de $y + z = 2$ y $2y + 2z = 4$ es $x_p = (1, 1)$. El espacio nulo de A en la figura 2.2 contiene a $(-1, 1)$ y a todos sus múltiplos $x_n = (-c, c)$:

Solución completa $\quad \begin{matrix} y + z = 2 \\ 2y + 2z = 4 \end{matrix}$ es resuelto por $x_p + x_n = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 - c \\ 1 + c \end{bmatrix}.$

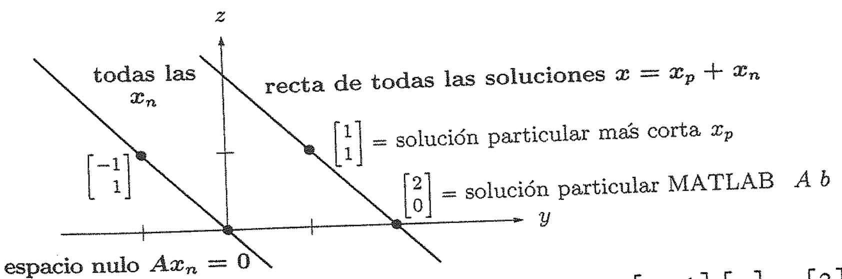


Figura 2.2 Las rectas paralelas de soluciones de $Ax_n = 0$ y $\begin{bmatrix}1 & 1\\ 2 & 2\end{bmatrix}\begin{bmatrix}y\\ z\end{bmatrix} = \begin{bmatrix}2\\ 4\end{bmatrix}$.

Forma escalonada *U* y forma reducida por renglones *R*

Se empieza por simplificar la matriz de 3 por 4, primero a U y luego a R:

$$\textbf{Ejemplo básico} \quad A = \begin{bmatrix} 1 & 3 & 3 & 2\\ 2 & 6 & 9 & 7\\ -1 & -3 & 3 & 4 \end{bmatrix}.$$

El pivote $a_{11} = 1$ es diferente de cero. Las operaciones elementales de costumbre producen ceros en la primera columna abajo de este pivote. Las malas noticias aparecen en la columna 2:

$$\textbf{En la columna 2 no hay pivote} \quad A \to \begin{bmatrix} 1 & 3 & 3 & 2\\ 0 & \mathbf{0} & 3 & 3\\ 0 & \mathbf{0} & 6 & 6 \end{bmatrix}.$$

El candidato a segundo pivote se volvió cero: *esto es inaceptable*. Abajo de ese cero se busca un elemento diferente de cero, intentando realizar un intercambio de renglones. En este caso *el elemento en cuestión también es cero*. Si A fuese cuadrada, lo anterior indicaría que la matriz es singular. Con una matriz singular, de cualquier forma es de esperar problemas, y no hay ninguna razón para detenerse. Todo lo que puede hacerse es *proceder a la siguiente columna*, donde el elemento pivote es 3. Al restar dos veces el segundo renglón del tercero, se llega a U:

$$\textbf{Matriz escalonada } U \quad U = \begin{bmatrix} \mathbf{1} & 3 & 3 & 2\\ 0 & 0 & \mathbf{3} & 3\\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Hablando estrictamente, se procede a la cuarta columna. En la tercera posición pivote hay un cero, y no puede hacerse nada. U es triangular superior, pero sus pivotes no están en la diagonal principal. Los elementos diferentes de cero de U muestran un "patrón escalonado", o ***forma escalonada***. Para el caso de 5 por 8 en la figura 2.3, los elementos indicados con un asterisco pueden o no ser cero.

Esta forma escalonada U, siempre puede obtenerse, con cero abajo de los pivotes:

1. Los pivotes son los primeros elementos diferentes de cero en sus renglones.
2. Abajo de cada pivote hay una columna de ceros, obtenida por eliminación.
3. Cada pivote está a la derecha del pivote en el renglón de arriba. Esto produce el patrón escalonado, y los renglones cero aparecen al último.

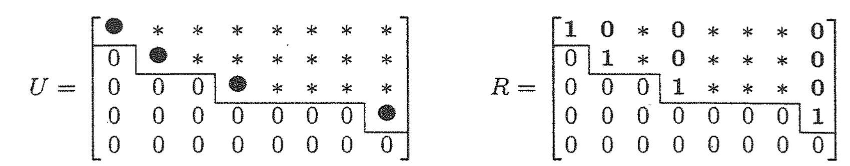


Figura 2.3 Elementos de una matriz escalonada U de 5 por 8 y su forma reducida R.

Debido a que se empezó con A y se terminó con U, seguramente el lector preguntará: ¿Se tiene $A = LU$ como antes? No hay razón para lo contrario, ya que los pasos de eliminación no cambiaron. Cada paso sigue restando un múltiplo de un renglón de un renglón abajo de él. El inverso de cada paso, suma el múltiplo que fue restado. Estos inversos vienen en el orden correcto para colocar los multiplicadores directamente en L:

$$\textbf{Triangular inferior} \quad L = \begin{bmatrix} 1 & 0 & 0\\ 2 & 1 & 0\\ -1 & 2 & 1 \end{bmatrix} \quad \text{y} \quad A = LU.$$

Observe que L es cuadrada. Tiene el mismo número de renglones que A y U.

La única operación que no es necesaria en nuestro ejemplo, aunque sí es necesaria en términos generales, es el intercambio de renglones por una matriz permutación P. Debido a que cuando no hay pivotes disponibles se prosigue con la siguiente columna, no es necesario suponer que A es no singular. A continuación se proporciona $PA = LU$ para todas las matrices:

2B Para cualquier matriz A de m por n existe una permutación P, una matriz triangular inferior L con diagonal unitaria, y una matriz U escalonada de m por n, tales que $PA = LU$.

Ahora viene R. Es posible avanzar más que hasta U, con la finalidad de hacer aún más sencilla a la matriz. El segundo renglón se divide entre su pivote 3, de modo que ***todos los pivotes son* 1**. Luego, el renglón pivote se utiliza para obtener ***cero arriba del pivote***. Esta vez, se resta un renglón de *un renglón más arriba*. El resultado final (la mejor forma que puede obtenerse) es la ***forma escalonada por renglones R***:

$$\begin{bmatrix} 1 & 3 & 3 & 2\\ 0 & 0 & 3 & 3\\ 0 & 0 & 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 3 & 3 & 2\\ 0 & 0 & \mathbf{1} & \mathbf{1}\\ 0 & 0 & 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{1} & \mathbf{3} & \mathbf{0} & \mathbf{-1}\\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{1}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} = R.$$

La matriz R es el resultado final de la eliminación sobre A. MATLAB usa el comando $R =$ rref(A). ¡Por supuesto, rref(R) proporciona nuevamente R!

¿Cuál es la forma reducida por renglones de una matriz invertible cuadrada? En este caso R es la *matriz identidad*. Hay un conjunto completo de pivotes, todos iguales a 1, con ceros arriba y abajo. Así, rref(A) $= I$, cuando A es invertible.

Para una matriz de 5 por 8 con cuatro pivotes, en la figura 2.3 se muestra la forma reducida R. **Sigue conteniendo una matriz identidad en los cuatro renglones pivote, y en las cuatro columnas pivote.** A partir de R rápidamente se encuentra el espacio nulo de A. $Rx = 0$ tiene las mismas soluciones que $Ux = 0$ y $Ax = 0$.

Variables pivote y variables libres

Nuestra meta es leer todas las soluciones de $Rx = 0$. Los pivotes son cruciales:

Espacio nulo de R (las columnas pivote están en negritas)

$$Rx = \begin{bmatrix} \mathbf{1} & 3 & \mathbf{0} & -1 \\ \mathbf{0} & 0 & \mathbf{1} & 1 \\ \mathbf{0} & 0 & \mathbf{0} & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Las incógnitas u, v, w, y se presentan por grupos. Un grupo contiene las ***variables pivote***, las que corresponden a las ***columnas con pivotes***. Las columnas primera y tercera contienen a los pivotes, de modo que u y w son las variables pivote. El otro grupo está integrado por las ***variables libres***, que corresponden a las ***columnas sin pivotes***. Éstas son las columnas segunda y cuarta, por lo que las variables libres son v y y.

Para encontrar la solución más general de $Rx = 0$ (o, de manera equivalente, de $Ax = 0$), es posible asignar valores arbitrarios a las variables libres. Suponga que estos valores se denominan simplemente v y y. Las variables pivote están determinadas completamente en términos de v y y:

$$\boldsymbol{Rx = 0} \qquad \begin{matrix} u + 3v - y = 0 & \text{produce} & u = -3v + y \\ w + y = 0 & \text{produce} & w = \quad\; -y \end{matrix} \tag{1}$$

Hay un "doble infinito" de soluciones, con v y y libres e independientes. La solución completa es una combinación de dos **soluciones especiales**:

El espacio nulo contiene a todas las combinaciones de las soluciones especiales

$$x = \begin{bmatrix} -3v + y \\ v \\ -y \\ y \end{bmatrix} = v \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}. \tag{2}$$

Por favor, observe nuevamente esta solución completa de $Rx = 0$ y $Ax = 0$. La solución especial $(-3, 1, 0, 0)$ tiene a las variables libres $v = 1$, y $y = 0$. Las variables libres de la otra solución especial $(1, 0, -1, 1)$ son $v = 0$, y $y = 1$. *Todas las soluciones son combinaciones lineales de estas dos soluciones.* La mejor forma de encontrar todas las soluciones de $Ax = 0$ es a partir de las soluciones especiales:

1. Después de llegar a $Rx = 0$, identificar las variables pivote, y las variables libres.
2. Asignar a una variable libre el valor 1, igualar a cero las demás variables libres, y resolver $Rx = 0$ para las variables pivote. Esta x es una solución especial.
3. Cada variable libre produce su propia "solución especial" por el paso 2. Las combinaciones de las soluciones especiales constituyen el espacio nulo: todas las soluciones de $Ax = 0$.

Dentro del espacio tetradimensional de todos los vectores posibles x, las soluciones de $Ax = 0$ constituyen un ***subespacio bidimensional***: el espacio nulo de A. En el ejemplo, $N(A)$ es generado por los vectores especiales $(-3, 1, 0, 0)$ y $(1, 0, -1, 1)$. Las combinaciones de estos dos vectores producen todo el espacio nulo.

Aquí hay un pequeño truco. Las soluciones especiales son especialmente fáciles a partir de R. Los números 3 y 0 y -1 y 1 están en las "columnas no pivote" de R. **Para encontrar las variables pivote** (no las libres)**, invierta sus signos en las soluciones especiales.** Escribiré las dos soluciones especiales de la ecuación (2) en una matriz N, de espacio nulo, para que el lector vea el patrón claro:

Matriz del espacio nulo (las columnas son soluciones especiales)

$$N = \begin{bmatrix} \mathbf{-3} & \mathbf{1} \\ 1 & 0 \\ \mathbf{0} & \mathbf{-1} \\ 0 & 1 \end{bmatrix} \begin{matrix} \text{no es libre} \\ \text{es libre} \\ \text{no es libre} \\ \text{es libre} \end{matrix}$$

Los valores de las variables libres son 1 y 0. Cuando las columnas libres se movieron al miembro derecho de la ecuación (2), sus coeficientes de 3 y 0 y -1 y 1 cambiaron de signo. Esto determinó las variables pivote en las soluciones especiales (las columnas de N).

Este es el momento de reconocer un teorema extremadamente importante. Suponga que una matriz tiene más columnas que renglones, $n > m$. Debido a que m renglones pueden aceptar cuando mucho a m pivotes, ***por lo menos debe haber $n - m$ variables libres***. Hay aún *más* variables libres si algunos renglones de R se reducen a cero; pero no importa el caso: por lo menos una variable debe ser libre. A esta variable libre puede asignarse cualquier valor, lo cual lleva a la siguiente conclusión:

2C Si $Ax = 0$ tiene más incógnitas que ecuaciones ($n > m$), por lo menos tiene una solución especial: hay más soluciones que la trivial $x = 0$.

Debe haber una infinidad de soluciones, ya que cualquier múltiplo cx también satisface $A(cx) = 0$. El espacio nulo contiene a la recta que pasa por x. Y si hay variables libres adicionales, el espacio nulo se convierte en más que justo una recta en el espacio n-dimensional. *El espacio nulo tiene la misma "dimensión" que el número de variables libres, y las soluciones especiales.*

La idea central: la ***dimensión*** de un subespacio, se precisará en la siguiente sección. Para el espacio nulo se cuentan las variables libres. ¡Para el espacio columna se cuentan las variables pivote!

Cómo resolver $Ax = b$, $Ux = c$ y $Rx = d$

El caso $b \neq 0$ es bastante distinto al caso $b = 0$. Las operaciones en los renglones sobre A también deben actuar sobre el miembro derecho (en b). Se empieza con las letras (b_1, b_2, b_3) para encontrar la condición de solubilidad para que b esté en el espacio columna. Luego se escoge $b = (1, 5, 5)$ y se encuentran todas las soluciones x.

Para el ejemplo original $Ax = b = (b_1, b_2, b_3)$, a ambos lados se aplican las operaciones que llevaron de A a U. El resultado es un sistema triangular superior $Ux = c$:

$$\boldsymbol{Ux = c} \qquad \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ y \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 - 2b_1 \\ b_3 - 2b_2 + 5b_1 \end{bmatrix}. \tag{3}$$

El vector c en el miembro derecho, que aparecía después de los pasos de eliminación hacia adelante, es justo $L^{-1}b$, como en el capítulo previo. Ahora se inicia con $Ux = c$.

No resulta evidente que estas ecuaciones tengan una solución. La tercera ecuación presenta serias dudas, ya que su miembro izquierdo es cero. ***Las ecuaciones son inconsistentes a menos que $b_3 - 2b_2 + 5b_1 = 0$***. Aunque haya más incógnitas que ecuaciones, puede no haber solución. Conocemos otra forma para contestar la misma pregunta: $Ax = b$ puede resolverse si y sólo si b está en el espacio columna de A. Este subespacio proviene de las cuatro columnas de A (¡no de U!):

Las columnas de A "generan" el espacio columna $\begin{bmatrix}1\\2\\-1\end{bmatrix}, \begin{bmatrix}3\\6\\-3\end{bmatrix}, \begin{bmatrix}3\\9\\3\end{bmatrix}, \begin{bmatrix}2\\7\\4\end{bmatrix}.$

Aunque hay cuatro vectores, sus combinaciones sólo llenan un plano en el espacio tridimensional. La columna 2 es tres veces la columna 1. La cuarta columna es igual a la tercera menos la primera. *Estas columnas dependientes, la segunda y la cuarta, son exactamente las que carecen de pivotes.*

El espacio columna $C(A)$ puede describirse en dos formas distintas. Por una parte, *es el plano generado por las columnas 1 y 3.* Las otras columnas están en ese plano, y no contribuyen con nada nuevo. De manera equivalente, es el plano de todos los vectores b que cumplen $b_3 - 2b_2 + 5b_1 = 0$; esta es la restricción si el sistema ha de ser resoluble. ***¡Toda columna satisface esta restricción, de modo que está forzada sobre b!*** Geométricamente, se verá que el vector $(5, -2, 1)$ es perpendicular a cada columna.

Si b pertenece al espacio columna, las soluciones de $Ax = b$ son fáciles de encontrar. La última ecuación en $Ux = c$ es $0 = 0$. A las variables libres v y y, es posible asignarles valores cualesquiera, como antes. Las variables pivote u y w siguen determinándose por sustitución hacia atrás. Para un ejemplo específico con $b_3 - 2b_2 + 5b_1 = 0$, se escoge $b = (1, 5, 5)$:

$$Ax = b \qquad \begin{bmatrix}1 & 3 & 3 & 2\\ 2 & 6 & 9 & 7\\ -1 & -3 & 3 & 4\end{bmatrix}\begin{bmatrix}u\\v\\w\\y\end{bmatrix} = \begin{bmatrix}1\\5\\5\end{bmatrix}.$$

La eliminación hacia adelante produce U a la izquierda y c a la derecha:

$$Ux = c \qquad \begin{bmatrix}1 & 3 & 3 & 2\\ 0 & 0 & 3 & 3\\ 0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}u\\v\\w\\y\end{bmatrix} = \begin{bmatrix}1\\3\\0\end{bmatrix}.$$

La última ecuación es $0 = 0$, como era de esperarse. La sustitución hacia atrás proporciona

$$\begin{aligned} 3w + 3y &= 3 \quad &\text{o} \quad w &= 1 - y \\ u + 3v + 3w + 2y &= 1 \quad &\text{o} \quad u &= -2 - 3v + y. \end{aligned}$$

De nuevo, hay una infinidad de soluciones: v y y son libres; u y w no lo son:

$$\textbf{Solución completa} \quad x = x_p + x_n \qquad x = \begin{bmatrix}u\\v\\w\\y\end{bmatrix} = \begin{bmatrix}-2\\0\\1\\0\end{bmatrix} + v\begin{bmatrix}-3\\1\\0\\0\end{bmatrix} + y\begin{bmatrix}1\\0\\-1\\1\end{bmatrix}. \tag{4}$$

Lo anterior contiene todas las soluciones de $Ax = 0$, más la nueva $x_p = (-2, 0, 1, 0)$. Esta x_p es una ***solución particular*** de $Ax = b$. Los dos últimos términos con v y y producen más soluciones (ya que cumplen $Ax = 0$). ***Toda solución de $Ax = b$ es la suma de una solución particular, y una solución de $Ax = 0$:***

$$x_{\text{completa}} = x_{\text{particular}} + x_{\text{espacio nulo}}$$

La solución particular en la ecuación (4) proviene de haber resuelto la ecuación *con todas las variables libres iguales a cero*. Esta es la única parte nueva, porque el espacio nulo ya estaba calculado. Cuando la ecuación en el recuadro se multiplica por A, se obtiene $Ax_{\text{completa}} = b + 0$.

Geométricamente, las soluciones llenan de nuevo una superficie bidimensional, pero no es un subespacio. No contiene a $x = 0$. *Es paralelo* al espacio nulo que se tenía anteriormente, desplazado por una solución particular x_p, como se muestra en la figura 2.2. La ecuación (4) es una forma aceptable de escribir la respuesta:

1. Reducir $Ax = b$ a $Ux = c$.
2. Con las variables libres $= 0$, encontrar una solución particular de $Ax_p = b$ y $Ux_p = c$.
3. Encontrar las soluciones especiales de $Ax = 0$ (o de $Ux = 0$ o de $Rx = 0$). Cada variable libre, a su vez, es 1. Así, $x = x_p +$ (cualquier combinación x_n de soluciones especiales).

Cuando la ecuación era $Ax = 0$, ¡la solución particular era el vector cero! Esto se ajusta al patrón, aunque $x_{\text{particular}} = 0$ no se escribió en la ecuación (2). Ahora x_p se sumó a las soluciones del espacio nulo, como en la ecuación (4).

Pregunta: ¿cómo la forma reducida R hace aún más clara esta solución? Se verá en el ejemplo. La ecuación 2 se resta de la ecuación 1, y luego la ecuación 2 se divide entre su pivote. En el miembro izquierdo, esto produce R, como antes. En el miembro derecho, estas operaciones cambian $c = (1, 3, 0)$ a un nuevo vector $d = (-2, 1, 0)$:

$$\textbf{Ecuación reducida } Rx = d \qquad \begin{bmatrix}1 & 3 & 0 & -1\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}u\\v\\w\\y\end{bmatrix} = \begin{bmatrix}-2\\1\\0\end{bmatrix}. \tag{5}$$

La solución particular x_p (una de muchas opciones) tiene las variables libres $v = y = 0$. Las columnas 2 y 4 pueden ignorarse. Luego, de inmediato se tiene $u = -2$ y $w = 1$, exactamente como en la ecuación (4). **Los elementos de d van directamente hacia x_p.** ¡Esto se debe a que la matriz identidad se encuentra en las columnas pivote de R!

A continuación se resumirá esta sección, antes de presentar otro ejemplo. La eliminación revela las variables pivote y las variables libres. ***Si hay r pivotes, entonces hay r variables pivote, y n − r variables libres.*** A este número importante r se asigna un nombre: se trata del ***rango de la matriz.***

2D Suponga que la eliminación reduce $Ax = b$ a $Ux = c$ y $Rx = d$, con r renglones pivote, y r columnas pivote. **El rango de estas matrices es r.** Los últimos $m - r$ renglones de U y R son cero, de modo que hay una solución sólo si los últimos $m - r$ elementos de c y d también son cero.

La solución completa es $x = x_p + x_n$. Una solución particular x_p tiene iguales a cero a todas las variables libres. Sus variables pivote son los primeros r elementos de d, por lo cual $Rx_p = d$.

Las soluciones del espacio nulo x_n son combinaciones de $n - r$ soluciones especiales, con una variable libre igual a 1. Las variables pivote en esa solución especial, pueden encontrarse en la columna correspondiente de R (con signo invertido).

Ahora el lector puede ver cómo es crucial el rango r. Éste cuenta los renglones pivote en el "espacio renglón" y las columnas pivote en el espacio columna. En el espacio nulo hay $n - r$ soluciones especiales. Hay $m - r$ condiciones de solubilidad sobre b o c o d.

Otro ejemplo desarrollado

La representación completa utiliza eliminación, y columnas pivote para encontrar el espacio columna, el espacio nulo y el rango. La matriz A de 3 por 4, tiene rango 2:

$$Ax = b \quad \text{es} \quad \begin{aligned} 1x_1 + 2x_2 + 3x_3 + \ \ 5x_4 &= b_1 \\ 2x_1 + 4x_2 + 8x_3 + 12x_4 &= b_2 \\ 3x_1 + 6x_2 + 7x_3 + 13x_4 &= b_3 \end{aligned} \qquad (6)$$

1. Reducir $[A\ b]$ a $[U\ c]$, para llegar a un sistema triangular $Ux = c$.
2. Encontrar la condición sobre b_1, b_2, b_3 para tener una solución.
3. Describir el espacio columna de A: ¿Cuál plano en $\mathbf{R}^3$?
4. Describir el espacio nulo de A: ¿Cuáles soluciones especiales en $\mathbf{R}^4$?
5. Encontrar una solución particular de $Ax = (0, 6, -6)$ y la solución completa $x_p + x_n$.
6. Reducir $[U\ c]$ a $[R\ d]$: soluciones especiales de R y x_p de d.

Solución (¡Observe cómo el miembro derecho se incluye como una columna adicional!)

1. Los multiplicadores en la eliminación son 2 y 3 y -1, tomando $[A\ b]$ a $[U\ c]$.

$$\left[\begin{array}{cccc|c} 1 & 2 & 3 & 5 & b_1 \\ 2 & 4 & 8 & 12 & b_2 \\ 3 & 6 & 7 & 13 & b_3 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & 2 & 3 & 5 & b_1 \\ 0 & 0 & 2 & 2 & b_2 - 2b_1 \\ 0 & 0 & -2 & -2 & b_3 - 3b_1 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & 2 & 3 & 5 & b_1 \\ 0 & 0 & 2 & 2 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - 5b_1 \end{array}\right].$$

2. La última ecuación muestra la condición de solubilidad $b_3 + b_2 - 5b_1 = 0$. Así, $0 = 0$.
3. El espacio columna de A es el plano que contiene a todas las combinaciones de las columnas pivote (1, 2, 3) y (3, 8, 7). **Segunda descripción:** El espacio columna contiene a todos los vectores con $b_3 + b_2 - 5b_1 = 0$. Esto hace resoluble a $Ax = b$, de modo que b está en el espacio columna. *Todas las columnas de A pasan esta prueba $b_3 + b_2 - 5b_1 = 0$. Esta es la ecuación del plano (en la primera descripción del espacio columna).*
4. Las soluciones especiales en N tienen las variables libres $x_2 = 1$, $x_4 = 0$ y $x_2 = 0$, $x_4 = 1$:

Matriz del espacio nulo
Soluciones especiales de $Ax = 0$
Sustitución hacia atrás en $Ux = 0$
Simplemente cambia signos en $Rx = 0$

$$N = \begin{bmatrix} -2 & -2 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{bmatrix}.$$

5. Se elige $b = (0, 6, -6)$, que tiene $b_3 + b_2 - 5b_1 = 0$. La eliminación lleva $Ax = b$ a $Ux = c = (0, 6, 0)$. Se sustituye hacia atrás con las variables libres $= 0$:

Solución particular de $Ax_p = (0, 6, -6)$

$$x_p = \begin{bmatrix} -9 \\ 0 \\ 3 \\ 0 \end{bmatrix} \begin{matrix} \\ \text{es libre} \\ \\ \text{es libre} \end{matrix}$$

La solución completa de $Ax = (0, 6, -6)$ es (esta x_p) + todas las x_n).

6. En la R reducida, la tercera columna cambia de (3, 2, 0) a (0, 1, 0). El miembro derecho $c = (0, 6, 0)$ se convierte en $d = (-9, 3, 0)$. Así, -9 y 3 van hacia x_p:

$$\begin{bmatrix} U & c \end{bmatrix} = \left[\begin{array}{cccc|c} 1 & 2 & 3 & 5 & 0 \\ 0 & 0 & 2 & 2 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \longrightarrow \begin{bmatrix} R & d \end{bmatrix} = \left[\begin{array}{cccc|c} \mathbf{1} & 2 & 0 & 2 & -9 \\ 0 & 0 & \mathbf{1} & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Esa matriz final $[R\ d]$ es rref($[A\ b]$) = rref($[U\ c]$). Los números 2 y 0 y 2 y 1 en las columnas libres de R tienen signo opuesto en las soluciones especiales (la matriz espacio nulo N). $Rx = d$, revela todo.

Conjunto de problemas 2.2

1. Construya un sistema con más incógnitas que ecuaciones, pero sin solución. Cambie el miembro derecho a cero, y encuentre todas las soluciones x_n.
2. Reduzca A y B a forma escalonada para encontrar sus rangos. ¿Qué variables son libres?

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

Encuentre las soluciones especiales de $Ax = 0$ y $Bx = 0$. Encuentre todas las soluciones.

3. Encuentre la forma escalonada de U, las variables libres, y las soluciones especiales:

$$A = \begin{bmatrix} 0 & 1 & 0 & 3 \\ 0 & 2 & 0 & 6 \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

$Ax = b$ es consistente (tiene una solución) cuando b satisface $b_2 =$ ____. Encuentre la solución completa en la misma forma que la ecuación (4).

4. Efectúe los mismos pasos que en el problema previo para encontrar la solución completa de $Mx = b$:

$$M = \begin{bmatrix} 0 & 0 \\ 1 & 2 \\ 0 & 0 \\ 3 & 6 \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}.$$

5. Escriba las soluciones completas $x = x_p + x_n$ de estos sistemas, como en la ecuación (4):

$$\begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 5 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

6. Describa el conjunto de miembros derechos b obtenibles (en el espacio columna) para

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},$$

al encontrar las restricciones sobre b que hacen que la tercera ecuación sea $0 = 0$ (después de la eliminación). ¿Cuál es el rango, y una solución particular?

7. Encuentre el valor de c que hace posible resolver $Ax = b$, y resuélvalo:

$$\begin{aligned} u + \ \ v + 2w &= 2 \\ 2u + 3v - \ \ w &= 5 \\ 3u + 4v + \ \ w &= c. \end{aligned}$$

8. ¿Cuáles deben ser las condiciones sobre b_1 y b_2 (en caso de haber alguna) para que $Ax = b$ tenga solución?

$$A = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 2 & 4 & 0 & 7 \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

Encuentre dos vectores en el espacio nulo de A, así como la solución completa de $Ax = b$.

9. *a*) Encuentre las soluciones especiales de $Ux = 0$. Reduzca U a R, y repita lo anterior:

$$Ux = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

b) Si el miembro derecho se cambia de $(0, 0, 0)$ a $(a, b, 0)$, ¿cuáles son todas las soluciones?

10. Encuentre un sistema $Ax = b$ de 2 por 3 cuya solución completa es

$$x = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + w \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}.$$

Encuentre un sistema de 3 por 3 con estas soluciones exactamente cuando $b_1 + b_2 = b_3$.

11. Escriba un sistema $Ax = b$ de 2 por 2 con muchas soluciones x_n, pero sin solución x_p. (En consecuencia, el sistema no tiene solución.) ¿Cuáles bs permiten una x_p?

12. ¿Cuáles de las siguientes reglas proporcionan una definición correcta del *rango* de A?

a) El número de renglones diferentes de cero en R.
b) El número de columnas menos el número total de renglones.
c) El número de columnas menos el número de columnas libres.
d) El número de 1s en R.

13. Encuentre las formas escalonadas reducidas por renglón R, y el rango de las siguientes matrices:

a) La matriz de 3 por 4 de todos los 1s.
b) La matriz de 4 por 4 con $a_{ij} = (-1)^{ij}$.
c) La matriz de 3 por 4 con $a_{ij} = (-1)^j$.

14. Encuentre R para cada una de las siguientes matrices (en bloque), así como las soluciones especiales:

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 3 \\ 2 & 4 & 6 \end{bmatrix} \qquad B = \begin{bmatrix} A & A \end{bmatrix} \qquad C = \begin{bmatrix} A & A \\ A & 0 \end{bmatrix}.$$

15. Si primero aparecen las r primeras variables pivote, entonces la R reducida debe verse como

$$R = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix} \qquad \begin{matrix} I \text{ es de } r \text{ por } r \\ F \text{ es de } r \text{ por } n - r \end{matrix}$$

¿Cuál es la matriz espacio nulo N que contiene las soluciones especiales?

16. Suponga que todas las r variables pivote aparecen *al último*. Describa los cuatro bloques en la forma escalonada reducida de m por n (el bloque B debe ser de r por r):

$$R = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

¿Cuál es la matriz espacio nulo N de las soluciones especiales? ¿Cuál es su forma?

17. (Problema tonto) Describa todas las matrices A_1 y A_2 de 2 por 3 con formas escalonadas por renglones R_1 y R_2, tales que $R_1 + R_2$ es la forma escalonada por renglones de $A_1 + A_2$. ¿Es cierto que $R_1 = A_1$ y $R_2 = A_2$ en este caso?

18. Si A tiene r columnas pivote, entonces A^T tiene r columnas pivote. Proporcione un ejemplo de 3 por 3 para el que los números de columna son distintos para A y A^T.

19. ¿Cuáles son las soluciones especiales de $Rx = 0$ y $R^Ty = 0$ para las siguientes R?

$$R = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 4 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad R = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

20. *Si el rango de A es r, entonces tiene una submatriz S de r por r que es invertible.* Encuentre esa submatriz S a partir de los renglones pivote, y las columnas pivote para cada A:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \end{bmatrix} \qquad A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} \qquad A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

21. Explique por qué los renglones pivote y las columnas pivote de A (no R) siempre proporcionan una submatriz invertible de A de r por r.

22. Encuentre los rangos de AB y AM (matriz de rango 1 multiplicada por una matriz de rango 1):

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 2 & 1 & 4 \\ 3 & 1.5 & 6 \end{bmatrix} \quad \text{y} \quad M = \begin{bmatrix} 1 & b \\ c & bc \end{bmatrix}.$$

23. Al multiplicar las matrices de rango 1 $A = uv^T$ y $B = wz^T$ se obtiene uz^T multiplicada por el número ____. El rango de AB es 1 a menos que ____ $= 0$.

24. Toda columna de AB es una combinación de las columnas de A. Así, las dimensiones de los espacios columna proporcionan **rango(AB) ≤ rango(A)**. Problema: También demuestre que **rango(AB) ≤ rango(B)**.

25. (Importante) Suponga que A y B son matrices de n por n, y que $AB = I$. A partir de rango(AB) ≤ rango(A), demuestre que el rango de A es n. Así, A es invertible, y B debe ser su inversa por ambos lados. En consecuencia, $BA = I$ (*¡lo cual no es tan evidente!*).

26. Si A es de 2 por 3 y C es de 3 por 2, a partir de su rango demuestre que $CA \neq I$. Proporcione un ejemplo en el que $AC = I$. Para $m < n$, una inversa derecha no es una inversa izquierda.

27. Suponga que A y B tienen *la misma* forma escalonada reducida por renglones R. Explique cómo cambiar de A a B, mediante operaciones elementales en los renglones. Así, B es igual a una matriz ____ multiplicada por A.

28. Toda matriz de m por n de rango r se reduce a (m por r) multiplicada por (r por n):

A = (columnas pivote de A)(r primeros renglones de R) = **(COL)(RENGLÓN)**.

Escriba la matriz de 3 por 4 A, al inicio de esta sección, como el producto de la matriz de 3 por 2 a partir de las columnas pivote, y la matriz de 2 por 4 de R:

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix}$$

29. Suponga que A es una matriz de m por n de rango r. Su forma escalonada reducida es R. Describa exactamente la *forma escalonada reducida por renglones de* R^T (no A^T).

30. (Recomendado) Ejecute los seis pasos que están después de la ecuación (6) para encontrar el espacio columna, y el espacio nulo de A y la solución de $Ax = b$:

$$A = \begin{bmatrix} 2 & 4 & 6 & 4 \\ 2 & 5 & 7 & 6 \\ 2 & 3 & 5 & 2 \end{bmatrix} \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix}.$$

31. Para cada c, encuentre R, y las soluciones especiales de $Ax = 0$:

$$A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 2 & 4 & 4 \\ 1 & c & 2 & 2 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 1-c & 2 \\ 0 & 2-c \end{bmatrix}.$$

32. ¿Cuál es la matriz espacio nulo N (de soluciones especiales) para A, B, C?

$$A = \begin{bmatrix} I & I \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} I & I \\ 0 & 0 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} I & I & I \end{bmatrix}.$$

Los problemas 33 a 36 son sobre la solución de $Ax = b$. Siga los pasos proporcionados en el texto para encontrar x_p y x_n. Reduzca la matriz aumentada $[A\ b]$.

33. Encuentre las soluciones completas de

$$\begin{aligned} x + 3y + 3z &= 1 \\ 2x + 6y + 9z &= 5 \\ -x - 3y + 3z &= 5 \end{aligned} \quad \text{y} \quad \begin{bmatrix} 1 & 3 & 1 & 2 \\ 2 & 6 & 4 & 8 \\ 0 & 0 & 2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}.$$

34. ¿Cuál debe ser la condición sobre b_1, b_2, b_3, para que el sistema sea resoluble? Incluya a b como una cuarta columna en $[A\ b]$. Encuentre todas las soluciones cuando se cumple esa condición:

$$\begin{aligned} x + 2y - 2z &= b_1 \\ 2x + 5y - 4z &= b_2 \\ 4x + 9y - 8z &= b_3. \end{aligned}$$

35. ¿Cuáles son las condiciones sobre b_1, b_2, b_3, b_4 para que cada uno de los siguientes sistemas sea resoluble? *Despeje* x:

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 2 & 5 \\ 3 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 2 & 5 & 7 \\ 3 & 9 & 12 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}.$$

36. ¿Cuáles vectores (b_1, b_2, b_3) están en el espacio columna de A? ¿Cuáles combinaciones de los renglones de A proporcionan cero?

$$a)\ A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 6 & 3 \\ 0 & 2 & 5 \end{bmatrix} \qquad b)\ A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 2 & 4 & 8 \end{bmatrix}.$$

37. ¿Por qué no es posible que un sistema de 1 por 3 tenga $x_p = (2, 4, 0)$ y $x_n =$ cualquier múltiplo de $(1, 1, 1)$?

38. *a)* Si $Ax = b$ tiene dos soluciones x_1 y x_2, encuentre dos soluciones de $Ax = 0$.
b) Luego, encuentre otra solución de $Ax = b$.

39. Explique por qué las siguientes afirmaciones son falsas:
a) La solución completa es cualquier combinación lineal de x_p y x_n.
b) Un sistema $Ax = b$ tiene cuando mucho una solución particular.
c) La solución x_p con todas las variables libres iguales a cero es la solución más breve (longitud mínima $\|x\|$). (Encuentre un contraejemplo de 2 por 2).
d) Si A es invertible, entonces no hay solución x_n en el espacio nulo.

40. Suponga que la columna 5 de U no tiene pivote. Entonces x_5 es una variable ____. El vector cero (es) (no es) la única solución de $Ax = 0$. Si $Ax = b$ tiene una solución, entonces tiene ____ soluciones.

41. Si se conocen x_p (variables libres $= 0$) y todas las soluciones especiales de $Ax = b$, encuentre x_p y todas las soluciones especiales para los siguientes sistemas:

$$Ax = 2b \qquad \begin{bmatrix} A & A \end{bmatrix} \begin{bmatrix} x \\ X \end{bmatrix} = b \qquad \begin{bmatrix} A \\ A \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} b \\ b \end{bmatrix}.$$

42. Si $Ax = b$ tiene una infinidad de soluciones, ¿por qué es imposible que $Ax = B$ (nuevo miembro derecho) tenga una sola solución? ¿Es posible que $Ax = B$ no tenga solución?

43. Escoja el número q de modo que (de ser posible) los rangos sean *a)* 1, *b)* 2, *c)* 3:

$$A = \begin{bmatrix} 6 & 4 & 2 \\ -3 & -2 & -1 \\ 9 & 6 & q \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 3 & 1 & 3 \\ q & 2 & q \end{bmatrix}.$$

44. Proporcione ejemplos de matrices A para las cuales el número de soluciones de $Ax = b$ sea
a) 0 o 1, dependiendo de b.
b) ∞, sin importar b.
c) 0 o ∞, dependiendo de b.
d) 1, sin importar b.

45. Escriba todas las relaciones conocidas entre r y m y n si $Ax = b$
a) No tiene solución para alguna b.
b) Tiene una infinidad de soluciones para toda b.
c) Tiene exactamente una solución para alguna b, y ninguna solución para otra b.
d) Tiene exactamente una solución para toda b.

46. Aplique la eliminación de Gauss-Jordan (el miembro derecho se convierte en una columna extra) a $Ux = 0$ y $Ux = c$. Obtenga $Rx = 0$ y $Rx = d$:

$$\begin{bmatrix} U & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & \mathbf{0} \\ 0 & 0 & 4 & \mathbf{0} \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} U & c \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & \mathbf{5} \\ 0 & 0 & 4 & \mathbf{8} \end{bmatrix}.$$

Resuelva $Rx = 0$ para encontrar x_n (su variable libre es $x_2 = 1$). Resuelva $Rx = d$ para encontrar x_p (su variable libre es $x_2 = 0$).

47. Aplique eliminación con la columna extra para obtener $Rx = 0$ y $Rx = d$:

$$\begin{bmatrix} U & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 6 & \mathbf{0} \\ 0 & 0 & 2 & \mathbf{0} \\ 0 & 0 & 0 & \mathbf{0} \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} U & c \end{bmatrix} = \begin{bmatrix} 3 & 0 & 6 & \mathbf{9} \\ 0 & 0 & 2 & \mathbf{4} \\ 0 & 0 & 0 & \mathbf{5} \end{bmatrix}.$$

Resuelva $Rx = 0$ (variable libre = 1). ¿Cuáles son las soluciones de $Rx = d$?

48. Reduzca lo siguiente a $Ux = c$ (eliminación gaussiana), y luego a $Rx = d$:

$$Ax = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 1 & 3 & 2 & 0 \\ 2 & 0 & 4 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 10 \end{bmatrix} = b.$$

Encuentre una solución particular x_p y todas las soluciones x_n del espacio nulo.

49. Encuentre A y B con la propiedad mencionada, o bien, explique por qué no es posible.

a) La única solución de $Ax = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ es $x = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

b) La única solución de $Bx = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ es $x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

50. La solución completa de $Ax = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ es $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Encuentre A.

51. El espacio nulo de una matriz A de 3 por 4 es la recta que pasa por (2, 3, 1, 0).

a) ¿Cuál es el *rango* de A y la solución completa de $Ax = 0$?

b) ¿Cuál es la forma escalonada reducida exacta por renglones R de A?

52. Reduzca las siguientes matrices A y B a sus formas escalonadas ordinarias U:

$$a)\ A = \begin{bmatrix} 1 & 2 & 2 & 4 & 6 \\ 1 & 2 & 3 & 6 & 9 \\ 0 & 0 & 1 & 2 & 3 \end{bmatrix} \qquad b)\ B = \begin{bmatrix} 2 & 4 & 2 \\ 0 & 4 & 4 \\ 0 & 8 & 8 \end{bmatrix}.$$

Encuentre una solución especial para cada variable libre, y describa cada solución de $Ax = 0$ y $Bx = 0$. Reduzca las formas escalonadas U a R, y destaque con un recuadro los renglones pivote y las columnas pivote de la matriz identidad.

53. ¿Falso o verdadero? (En caso de ser cierto, proporcione una razón, o un contraejemplo si es falso.)

a) Una matriz cuadrada no tiene variables libres.

b) Una matriz invertible no tiene variables libres.

c) Una matriz de m por n no tiene más de n variables pivote.

d) Una matriz de m por n no tiene más de m variables pivote.

54. ¿Existe una matriz de 3 por 3 sin elementos cero para la cual $U = R = I$?

55. Escriba tantos 1s como pueda en una matriz escalonada U de 4 por 7, y en una forma *reducida* R cuyas columnas pivote sean 2, 4, 5.

56. Suponga que la columna 4 de una matriz de 3 por 5 es toda igual a 0s. Entonces x_4 es ciertamente una variable ____. La solución especial para esta variable es el vector $x =$ ____.

57. Suponga que las columnas primera y última de una matriz de 3 por 5 son iguales (diferentes de cero). Entonces ____ es una variable libre. Encuentre las soluciones especiales para esa variable.

58. La ecuación $x - 3y - z = 0$ determina un plano en $\mathbf{R}^3$. ¿Cuál es la matriz A en esta ecuación? ¿Cuáles son las variables libres? Las soluciones especiales son (3, 1, 0) y ____. El plano paralelo $x - 3y - z = 12$ contiene al punto particular (12, 0, 0). Todos los puntos en este plano tienen la forma siguiente (escriba las primeras componentes):

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \ \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} \ \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} \ \\ 0 \\ 1 \end{bmatrix}.$$

59. Suponga que columna 1 + columna 3 + columna 5 = 0 en una matriz de 4 por 5 con cuatro pivotes. ¿Cuál columna carece seguramente de pivote? (¿Y cuál es la variable libre?) ¿Cuál es la solución especial? ¿Cuál es el espacio nulo?

En los problemas 60 a 66 se solicitan matrices (de ser posible) con ciertas propiedades.

60. Construya una matriz cuyo espacio nulo conste de todas las combinaciones de (2, 2, 1, 0) y (3, 1, 0, 1).

61. Construya una matriz cuyo espacio nulo conste de todos los múltiplos de (4, 3, 2, 1).

62. Construya una matriz cuyo espacio columna contenga a (1, 1, 5) y a (0, 3, 1), y cuyo espacio nulo contenga a (1, 1, 2).

63. Construya una matriz cuyo espacio columna contenga a (1, 1, 0) y a (0, 1, 1), y cuyo espacio nulo contenga a (1, 0, 1) y a (0, 0, 1).

64. Construya una matriz cuyo espacio columna contenga a (1, 1, 1) y cuyo espacio nulo es la recta de múltiplos de (1, 1, 1, 1).

65. Construya una matriz de 2 por 2 cuyo espacio nulo sea igual a su espacio columna.

66. ¿Por qué ninguna matriz de 3 por 3 tiene un espacio nulo igual a su espacio columna?

67. La forma reducida R de una matriz de 3 por 3 con elementos elegidos aleatoriamente casi seguramente es ____. ¿Cuál R es virtualmente cierta si la matriz aleatoria A es de 4 por 3?

68. Demuestre con un ejemplo que las tres siguientes afirmaciones suelen ser *falsas*:

a) A y A^{T} tienen el mismo espacio nulo.

b) A y A^{T} tienen las mismas variables libres.

c) Si R es la forma reducida de rref(A), entonces R^{T} es rref(A^{T}).

69. Si las soluciones especiales de $Rx = 0$ están en las columnas de estas N, vuelva hacia atrás para encontrar los renglones diferentes de cero de las matrices reducidas R:

$$N = \begin{bmatrix} 2 & 3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad N = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{y} \quad N = \begin{bmatrix} \ \end{bmatrix} \quad \text{(matriz vacía de 3 por 1).}$$

70. Explique por qué A y $-A$ siempre tienen la misma forma escalonada reducida R.

2.3 INDEPENDENCIA LINEAL, BASE Y DIMENSIÓN

Por sí mismos, los números m y n proporcionan una representación incompleta del verdadero tamaño de un sistema lineal. La matriz de nuestro ejemplo tenía tres renglones y cuatro columnas, aunque el tercer renglón era sólo una combinación de los dos primeros. Después de la eliminación se convirtió en un renglón cero. No afectó el problema homogéneo $Ax = 0$. Las cuatro columnas también fracasaron en cuanto a ser independientes, y el espacio columna degeneró en un plano bidimensional.

El número importante que está comenzando a surgir (el tamaño verdadero) es el ***rango r***. El rango se introdujo como *el número de pivotes* en el proceso de eliminación. De manera equivalente, la matriz final U tiene r renglones diferentes de cero. Esta definición hubiera podido proporcionarse a una computadora, aunque sería erróneo dejarla ahí, porque el rango posee un significado simple e intuitivo: *El rango cuenta el número de renglones genuinamente independientes en la matriz A*. Lo que se busca son definiciones matemáticas, más que computacionales.

El objetivo de esta sección es explicar y usar cuatro conceptos:

1. Independencia o dependencia lineal.
2. Generación de un subespacio.
3. Base de un subespacio (un conjunto de vectores).
4. Dimensión de un subespacio (un número).

El primer paso es definir ***independencia lineal***. Dado un conjunto de vectores $v_1, \ldots, v_k$, se buscan sus combinaciones $c_1v_1 + c_2v_2 + \cdots + c_kv_k$. La combinación trivial, con todos los pesos $c_i = 0$, evidentemente produce el vector cero: $0v_1 + \cdots + 0v_k = 0$. La pregunta es si ésta es la *única forma* de producir cero. En caso afirmativo, los vectores son independientes.

Si con cualquier otra combinación de los vectores se obtiene cero, entonces son *dependientes*.

> **2E** Suponga que $c_1v_1 + \cdots + c_kv_k = 0$ sólo ocurre cuando $c_1 = \cdots = c_k = 0$. Entonces los vectores $v_1, \ldots, v_k$ son ***linealmente independientes***. Si algunas cs son diferentes de cero, las vs son ***linealmente dependientes***. Un vector es una combinación de los demás.

La dependencia lineal es fácil de visualizar en el espacio tridimensional, cuando todos los vectores salen del origen. Dos vectores son dependientes si están en la misma recta. *Tres vectores son dependientes si están en el mismo plano*. Una elección aleatoria de tres vectores, sin ningún accidente especial, debe producir independencia lineal (no están en un plano). Cuatro vectores siempre son linealmente dependientes en $\mathbf{R}^3$.

Ejemplo 1 Si v_1 = vector cero, entonces el conjunto es linealmente dependiente. Puede elegirse $c_1 = 3$ y todas las demás $c_i = 0$; esta es una combinación no trivial que produce cero.

Ejemplo 2 Las columnas de la matriz

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix}$$

son linealmente dependientes, ya que la segunda columna es tres veces la primera. La combinación de las columnas con pesos $-3, 1, 0, 0$ proporciona una columna de ceros.

Los renglones también son linealmente dependientes; el renglón 3 es dos veces el renglón 2 menos cinco veces el renglón 1. (Esto es lo mismo que la combinación de b_1, b_2, b_3, que tuvo que desaparecerse en el miembro derecho para que $Ax = b$ fuera consistente. A menos que $b_3 - 2b_2 + 5b_1 = 0$, la tercera ecuación no se convierte en $0 = 0$).

Ejemplo 3 Las columnas de la siguiente matriz triangular son linealmente *independientes*:

$$\textbf{No hay ceros en la diagonal} \qquad A = \begin{bmatrix} 3 & 4 & 2 \\ 0 & 1 & 5 \\ 0 & 0 & 2 \end{bmatrix}.$$

Se busca una combinación de las columnas que produzca cero:

$$\textbf{Se resuelve } Ac = 0 \qquad c_1 \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 2 \\ 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Hay que demostrar que c_1, c_2, c_3 están obligados todos a ser cero. La última ecuación proporciona $c_3 = 0$. Luego, la siguiente ecuación proporciona $c_2 = 0$, y al sustituir en la primera ecuación se obliga a que $c_1 = 0$. La única combinación que produce el vector cero es la combinación trivial. ***El espacio nulo de A contiene sólo al vector cero $c_1 = c_2 = c_3 = 0$.***

> ***Las columnas de A son independientes exactamente cuando $N(A)$ = (vector cero).***

Un razonamiento semejante es válido para los renglones de A, que también son independientes. Suponga que

$$c_1(3, 4, 2) + c_2(0, 1, 5) + c_3(0, 0, 2) = (0, 0, 0).$$

A partir de las primeras componentes se encuentra que $3c_1 = 0$ o $c_1 = 0$. Luego, las segundas componentes proporcionan $c_2 = 0$, y finalmente $c_3 = 0$.

Los renglones diferentes de cero de una matriz en forma escalonada U deben ser independientes. Además, si se eligen *las columnas que contienen a los pivotes*, también son linealmente independientes. En el ejemplo previo, con

$$\begin{matrix}\textbf{Dos renglones independientes} \\ \textbf{Dos columnas independientes}\end{matrix} \qquad U = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

las columnas pivote 1 y 3 son independientes. Ningún conjunto de tres columnas es independiente, y ciertamente tampoco ninguno de cuatro. Es cierto que las columnas 1 y 4 también son independientes, pero si el último 1 cambia a 0, entonces serían dependientes. *Por tanto, lo que garantiza su independencia son las columnas que contienen a los pivotes.* A continuación se proporciona la regla general:

> **2F** Los r renglones diferentes de cero de una matriz escalonada U y una matriz reducida R son linealmente independientes. También las r columnas que contienen a los pivotes son linealmente independientes.

Ejemplo 4 Las columnas de la matriz identidad de n por n son independientes:

$$I = \begin{bmatrix} 1 & 0 & \cdot & 0 \\ 0 & 1 & \cdot & 0 \\ \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Estas columnas $e_1, \ldots, e_n$ representan vectores unitarios en las direcciones coordenadas; en $\mathbf{R}^4$,

$$e_1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \quad e_2 = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \quad e_3 = \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \quad e_4 = \begin{bmatrix}0\\0\\0\\1\end{bmatrix}.$$

La mayor parte de los conjuntos de cuatro vectores en $\mathbf{R}^4$ son independientes. Estos *es* podrían ser el conjunto más seguro.

Para comprobar la independencia de un conjunto de vectores $v_1, \ldots, v_n$, éstos deben escribirse en las columnas de A. Luego se resuelve el sistema $Ac = 0$; los vectores son dependientes si hay otra solución aparte de $c = 0$. Si no hay variables libres (*rango n*), no existe espacio nulo excepto $c = 0$; los vectores son independientes. Si el rango es menor que n, entonces por lo menos una variable libre puede ser diferente de cero y las columnas son dependientes.

Un caso reviste especial importancia. Considere que los n vectores tienen m componentes, de modo que A es una matriz de m por n. Ahora suponga que $n > m$. ¡Hay demasiadas columnas para ser independientes! No puede haber n pivotes, ya que no hay suficientes renglones para acomodarlos. El rango es menor que n. Todo sistema $Ac = 0$ con más incógnitas que ecuaciones tiene soluciones $c \neq 0$.

2G Un conjunto de n vectores en $\mathbf{R}^m$ debe ser linealmente dependiente si $n > m$.

El lector reconocerá esto como una forma disfrazada de 2C: Todo sistema $Ax = 0$ de m por n tiene soluciones diferentes de cero si $n > m$.

Ejemplo 5 Las tres siguientes columnas en $\mathbf{R}^2$ no pueden ser independientes:

$$A = \begin{bmatrix}1 & 2 & 1\\1 & 3 & 2\end{bmatrix}.$$

Para encontrar la combinación de las columnas con la que se obtiene cero se resuelve $Ac = 0$:

$$A \to U = \begin{bmatrix}1 & 2 & 1\\0 & 1 & 1\end{bmatrix}.$$

Si a la variable libre c_3 se asigna el valor 1, entonces la sustitución hacia atrás en $Uc = 0$ proporciona $c_2 = -1$, $c_1 = 1$. Con estos tres pesos, la primera columna menos la segunda más la tercera es igual a cero: Dependencia.

Cómo generar un subespacio

A continuación se definirá lo que se entiende por el hecho de que un conjunto de vectores *generen un espacio*. El espacio columna de A *es generado* por las columnas. **Su combinación produce todo el espacio:**

2H Si un espacio vectorial $\mathbf{V}$ consta de todas las combinaciones lineales de $w_1, \ldots, w_\ell$, entonces estos vectores ***generan*** el espacio. Todo vector v en $\mathbf{V}$ es alguna combinación de las ws:
Todo v proviene de ws $v = c_1w_1 + \cdots + c_\ell w_\ell$ para algunos coeficientes c_i.

Se permite que una combinación distinta de ws proporcione el mismo vector v. Las cs no necesitan ser únicas, porque el conjunto generador puede ser excesivamente grande: puede incluir el vector cero, e incluso a todos los vectores.

Ejemplo 6 Los vectores $w_1 = (1, 0, 0)$, $w_2 = (0, 1, 0)$, y $w_3 = (-2, 0, 0)$, generan un plano (el plano x-y) en $\mathbf{R}^3$. Los dos primeros vectores también generan este plano, mientras w_1 y w_3 sólo generan una recta.

Ejemplo 7 El espacio columna de A es exactamente ***el espacio generado por sus columnas***. El espacio renglón es generado por los renglones. La definición se hace para ordenar. Al multiplicar A por cualquier x se obtiene una combinación de las columnas; es un vector Ax en el espacio columna.

Los vectores de coordenadas $e_1, \ldots, e_n$ que provienen de la matriz identidad generan $\mathbf{R}^n$. Todo vector $b = (b_1, \ldots, b_n)$ es una combinación de estas columnas. En este ejemplo los pesos son las componentes b_i mismas: $b = b_1e_1 + \cdots + b_ne_n$. Sin embargo, ¡las columnas de otras matrices también generan $\mathbf{R}^n$!

Base de un espacio vectorial

Para decidir si b es una combinación de las columnas, se intenta resolver $Ax = b$. Para decidir si las columnas son independientes, se resuelve $Ax = 0$. ***La generación implica al espacio columna, y la independencia implica al espacio nulo.*** Los vectores de coordenadas $e_1, \ldots, e_n$ generan $\mathbf{R}^n$ y son linealmente independientes. En términos bastos, ***ningún vector en ese conjunto se desperdicia***. Esto conduce al concepto crucial de ***base***.

2I Una ***base*** de $\mathbf{V}$ es una sucesión de vectores que cumplen dos propiedades a la vez:

1. Los vectores son linealmente independientes (no hay demasiados vectores).
2. Los vectores generan el espacio $\mathbf{V}$ (no hay muy pocos vectores).

Esta combinación de propiedades es absolutamente fundamental para el álgebra lineal. Significa que todo vector en el espacio es una combinación de los vectores de la base, ya que éstos generan el espacio. También significa que la combinación es única: Si $v = a_1v_1 + \cdots + a_kv_k$ y también $v = b_1v_1 + \cdots + b_kv_k$, entonces al restar se obtiene $0 = \Sigma(a_i - b_i)v_i$. Es ahora que entra en juego la independencia: todo coeficiente $a_i - b_i$ debe ser cero. En consecuencia, $a_i = b_i$. ***Hay una y sólo una forma de escribir v como una combinación lineal de los vectores de la base.***

Hubiera sido mejor decir de una vez que los vectores de coordenadas $e_1, \ldots, e_n$ no son la única base de $\mathbf{R}^n$. Algunas cosas en álgebra son únicas, pero no ésta. Un espacio vectorial ***tiene una infinidad de bases distintas***. Siempre que una matriz cuadrada es invertible, sus columnas son independientes, de modo que son una base de $\mathbf{R}^n$. Las dos columnas de la siguiente matriz no singular constituyen una base de $\mathbf{R}^2$:

$$A = \begin{bmatrix}1 & 1\\2 & 3\end{bmatrix}$$

Todo vector bidimensional es una combinación de estas columnas (¡independientes!).

Ejemplo 8 El plano x-y de la figura 2.4, es justo $\mathbf{R}^2$. El vector v_1 en sí es linealmente independiente, pero fracasa en generar $\mathbf{R}^2$. Los tres vectores v_1, v_2, v_3 ciertamente generan $\mathbf{R}^2$, pero no son independientes. *Dos* vectores *cualesquiera* de esos, por ejemplo v_1 y v_2, tienen ambas propiedades: generan el espacio y son independientes. Así, constituyen una base. Observe de nuevo que *un espacio vectorial no tiene una base única.*

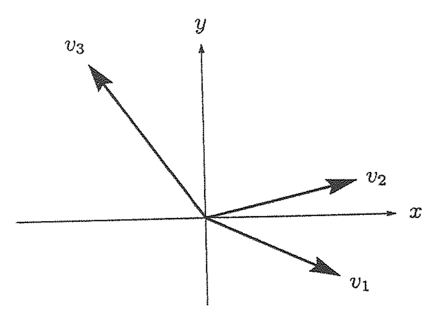


Figura 2.4 Conjunto generador v_1, v_2, v_3. Bases v_1, v_2 y v_1, v_3 y v_2, v_3.

Ejemplo 9 Las cuatro siguientes columnas generan el espacio columna U, pero no son independientes:

$$\textbf{Matriz escalonada} \quad U = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Hay muchas posibilidades para la base, pero se propone una opción específica: ***Las columnas que contienen pivotes*** (en este caso la primera y la tercera, que corresponden a las variables básicas) ***constituyen una base de espacio columna***. Estas columnas son independientes, y es fácil ver que generan el espacio. De hecho, el espacio columna de U es justo el plano x-y dentro de $\mathbf{R}^3$. $\mathbf{C}(U)$ *no es lo mismo* que el espacio columna $\mathbf{C}(A)$ antes de la eliminación; sin embargo, *el número* de columnas independientes no cambia.

En resumen: *Las columnas de cualquier matriz generan su espacio columna*. Si son independientes, constituyen una base del espacio columna, sin importar que la matriz sea cuadrada o rectangular. Si se pide que las columnas sean una base para todo el espacio $\mathbf{R}^n$, entonces la matriz debe ser cuadrada e invertible.

Dimensión de un espacio vectorial

Un espacio tiene una infinidad de bases distintas, aunque todas estas opciones tienen algo en común. El ***número de vectores de la base*** es una propiedad del espacio en sí:

> **2J** Dos bases cualesquiera de un espacio vectorial $\mathbf{V}$ contienen el mismo número de vectores. Este número que es compartido por todas las bases y expresa el número de "grados de libertad" del espacio, es la ***dimensión*** de $\mathbf{V}$.

Hay que demostrar este hecho: Todas las bases posibles contienen el mismo número de vectores. El plano x-y en la figura 2.4 tiene dos vectores en cada base; su dimensión es 2.

En tres dimensiones se requieren tres vectores, ya sea a lo largo de los ejes x-y-z ó en otras tres direcciones (¡linealmente independientes!) ***La dimensión del espacio $\mathbf{R}^n$ es n.*** La dimensión del espacio columna de U en el ejemplo 9 era 2; era un "subespacio bidimensional de $\mathbf{R}^3$." La matriz cero es más bien excepcional, ya que su espacio columna contiene sólo al vector cero. Por convencionalismo, el conjunto vacío es una base de ese espacio, y su dimensión es cero.

A continuación se presenta el primer gran teorema en álgebra lineal:

> **2K** Si $v_1, \ldots, v_m$ y $w_1, \ldots, w_n$ son dos bases del mismo espacio vectorial, entonces $m = n$. El número de vectores es el mismo.

Demostración Suponga que hay más ws que vs ($n > m$). Se llegará a una contradicción. Debido a que las vs constituyen una base, deben generar el espacio. *Toda w_j puede escribirse como una combinación de las vs*: si $w_1 = a_{11}v_1, + \cdots + a_{m1}v_m$, esta es la primera columna de una multiplicación de matrices VA:

$$W = \begin{bmatrix} w_1 & w_2 & \ldots & w_n \end{bmatrix} = \begin{bmatrix} v_1 & \ldots & v_m \end{bmatrix} \begin{bmatrix} a_{11} & & \\ \vdots & & \\ a_{m1} & & \end{bmatrix} = VA.$$

No se conoce cada a_{ij}, pero sí la forma de A (es de m por n). El segundo vector w_2 también es una combinación de las vs. Los coeficientes de esa combinación llenan la segunda columna de A. La clave es que A tiene un renglón para cada v y una columna para cada w. A es una matriz corta pero ancha, ya que $n > m$. ***Hay una solución diferente de cero para $Ax = \mathbf{0}$.*** Así, $VAx = 0$, que es $Wx = 0$. ¡*Una combinación de las ws da cero*! Las ws podrían no ser una base, de modo que no es posible que $n > m$.

Si $m > n$, se intercambian las vs y las ws y se repiten los mismos pasos. La única forma de evitar una contradicción es si $m = n$. Así se termina la demostración de que $m = n$. Repitiendo: ***La dimensión de un espacio*** es el número de vectores que hay en cada base. ■

Esta demostración se usó antes para demostrar que todo conjunto de $m + 1$ vectores en $\mathbf{R}^m$ debe ser dependiente. Las vs y las ws no necesitan ser vectores columna: la demostración fue toda concerniente a la matriz A de coeficientes. De hecho, puede verse este resultado general: *En un subespacio de dimensión k, ningún conjunto de más de k vectores puede ser independiente, y ningún conjunto con menos de k vectores puede generar el espacio.*

Hay otros teoremas "duales", de los cuales sólo se menciona uno. Puede empezarse con un conjunto de vectores que sea demasiado pequeño o demasiado grande, y terminar con la base siguiente:

> **2L** Cualquier conjunto linealmente independiente en $\mathbf{V}$ puede extenderse a una base, sumando más vectores si es necesario.
> Cualquier conjunto generado en $\mathbf{V}$ puede reducirse a una base, eliminando vectores si es necesario.

La cuestión es que una base es un ***conjunto independiente maximal***. No puede hacerse más grande sin perder independencia. Una base también es un ***conjunto independiente minimal***. No es posible hacerlo más pequeño y seguir generando el espacio.

El lector debe observar que el término "dimensional" se utiliza en dos sentidos diferentes. Se habla sobre un ***vector*** tetradimensional, haciendo referencia a un vector que está en

$\mathbf{R}^4$. Ahora se ha definido un ***subespacio*** tetradimensional; un ejemplo es el conjunto de vectores que están en $\mathbf{R}^6$ cuyas componentes primera y última son cero. Los elementos de este subespacio tetradimensional son vectores hexadimensionales como (0, 5, 1, 3, 4, 0).

Una nota final sobre el lenguaje que se utiliza en álgebra lineal. Nunca se utilizan expresiones como "la base de una matriz", "rango de un espacio" o "dimensión de una base". Estas frases carecen de sentido. *Es la dimensión del espacio columna* lo que es igual *al rango de una matriz*, como se demostrará en la siguiente sección.

Conjunto de problemas 2.3

Los problemas 1 a 10 son sobre independencia, y dependencia lineal.

1. Demestre que v_1, v_2, v_3 son independientes pero que v_1, v_2, v_3, v_4 son dependientes:

$$v_1 = \begin{bmatrix}1\\0\\0\end{bmatrix} \quad v_2 = \begin{bmatrix}1\\1\\0\end{bmatrix} \quad v_3 = \begin{bmatrix}1\\1\\1\end{bmatrix} \quad v_4 = \begin{bmatrix}2\\3\\4\end{bmatrix}.$$

Resuelva $c_1v_1 + \cdots + c_4v_4 = 0$ o bien $Ac = 0$. Las vs van en las columnas de A.

2. Encuentre el mayor número posible de vectores independientes entre los siguientes:

$$v_1 = \begin{bmatrix}1\\-1\\0\\0\end{bmatrix} \quad v_2 = \begin{bmatrix}1\\0\\-1\\0\end{bmatrix} \quad v_3 = \begin{bmatrix}1\\0\\0\\-1\end{bmatrix} \quad v_4 = \begin{bmatrix}0\\1\\-1\\0\end{bmatrix} \quad v_5 = \begin{bmatrix}0\\1\\0\\-1\end{bmatrix} \quad v_6 = \begin{bmatrix}0\\0\\1\\-1\end{bmatrix}.$$

Este número es la ____ del espacio generado por las vs.

3. Demuestre que si $a = 0$, $d = 0$ o bien $f = 0$ (3 casos), entonces las columnas de U son dependientes:

$$U = \begin{bmatrix}a & b & c\\0 & d & e\\0 & 0 & f\end{bmatrix}.$$

4. Si a, d, f en el problema 3 son todos distintos de cero, demuestre que la única solución de $Ux = 0$ es $x = 0$. Así, las columnas de U son independientes.

5. Decida la dependencia o independencia de
   a) Los vectores (1, 3, 2), (2, 1, 3), y (3, 2, 1).
   b) Los vectores (1, −3, 2), (2, 1, −3), y (−3, 2, 1).

6. Escoja tres columnas independientes de U. Luego elija otras dos. Haga lo mismo para A. ¿Encontró bases para estos espacios?

$$U = \begin{bmatrix}2 & 3 & 4 & 1\\0 & 6 & 7 & 0\\0 & 0 & 0 & 9\\0 & 0 & 0 & 0\end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix}2 & 3 & 4 & 1\\0 & 6 & 7 & 0\\0 & 0 & 0 & 9\\4 & 6 & 8 & 2\end{bmatrix}.$$

7. Si w_1, w_2, w_3 son vectores independientes, demuestre que las diferencias $v_1 = w_2 - w_3$, $v_2 = w_1 - w_3$, y $v_3 = w_1 - w_2$ son *dependientes*. Encuentre una combinación de las vs que dé cero.

8. Si w_1, w_2, w_3 son vectores independientes, demuestre que las sumas $v_1 = w_2 + w_3$, $v_2 = w_1 + w_3$, y $v_3 = w_1 + w_2$ son *independientes*. (Escriba $c_1v_1 + c_2v_2 + c_3v_3 = 0$ en términos de las ws. Encuentre, y resuelva ecuaciones para las cs.)

9. Suponga que v_1, v_2, v_3 y v_4 son vectores en $\mathbf{R}^3$.
   a) Estos cuatro vectores son dependientes porque ____.
   b) Los dos vectores v_1 y v_2 son dependientes porque ____.
   c) Los vectores v_1 y (0, 0, 0) son dependientes porque ____.

10. Encuentre dos vectores independientes en el plano $x + 2y - 3z - t = 0$ en $\mathbf{R}^4$. Luego encuentre tres vectores independientes. ¿Por qué no cuatro? ¿De cuál matriz es éste el espacio nulo?

Los problemas 11 a 18 son sobre el espacio *generado* por un conjunto de vectores. Considere todas las combinaciones lineales de los vectores.

11. Describa el subespacio de $\mathbf{R}^3$ (¿es una recta, un plano o $\mathbf{R}^3$?) generado por
    a) Los vectores (1, 1, −1) y (−1, −1, 1).
    b) Los vectores (0, 1, 1) y (1, 1, 0) y (0, 0, 0).
    c) Las columnas de una matriz escalonada de 3 por 5 con 2 pivotes.
    d) Todos los vectores con componentes positivas.

12. El vector b está en el subespacio generado por las columnas de A cuando hay una solución de ____. El vector c está en el espacio renglón de A cuando hay una solución de ____. ¿*Falso o verdadero*?: Si el vector cero está en el espacio renglón, los renglones son dependientes.

13. Encuentre las dimensiones de
    a) El espacio columna de A,
    b) El espacio columna de U,
    c) El espacio renglón de A,
    d) El espacio renglón de U.

    ¿Cuáles de estos dos espacios son iguales?

$$A = \begin{bmatrix}1 & 1 & 0\\1 & 3 & 1\\3 & 1 & -1\end{bmatrix} \quad \text{y} \quad U = \begin{bmatrix}1 & 1 & 0\\0 & 2 & 1\\0 & 0 & 0\end{bmatrix}.$$

14. Escoja $x = (x_1, x_2, x_3, x_4)$ en $\mathbf{R}^4$. Tiene 24 reordenamientos como (x_2, x_1, x_3, x_4) y (x_4, x_3, x_1, x_2). Estos 24 vectores, incluyendo x mismo, generan un subespacio $\mathbf{S}$. Encuentre vectores específicos x de modo que la dimensión de $\mathbf{S}$ sea: *a*) 0, *b*) 1, *c*) 3, *d*) 4.

15. $v + w$ y $v - w$ son combinaciones de v y w. Escriba v y w como combinaciones de $v + w$ y $v - w$. Los dos pares de vectores ____ el mismo espacio. ¿Cuándo constituyen una base para el mismo espacio?

16. Establezca si los siguientes vectores son o no linealmente independientes, resolviendo $c_1v_1 + c_2v_2 + c_3v_3 + c_4v_4 = 0$:

$$v_1 = \begin{bmatrix}1\\1\\0\\0\end{bmatrix}, \quad v_2 = \begin{bmatrix}1\\0\\1\\0\end{bmatrix}, \quad v_3 = \begin{bmatrix}0\\0\\1\\1\end{bmatrix}, \quad v_4 = \begin{bmatrix}0\\1\\0\\1\end{bmatrix}.$$

También decida si generan $\mathbf{R}^4$, intentando resolver $c_1v_1 + \cdots + c_4v_4 = (0, 0, 0, 1)$.

17. Suponga que los vectores cuya independencia habrá de comprobarse se escriben en los renglones, en vez de en las columnas de A. ¿Cómo decide el proceso de eliminación de A a U si los vectores son o no independientes?

18. Para decidir si b está en el subespacio generado por $w_1, \ldots, w_n$, considere que los vectores w son las columnas de A, e intente resolver $Ax = b$. ¿Cuál es el resultado para
   a) $w_1 = (1, 1, 0)$, $w_2 = (2, 2, 1)$, $w_3 = (0, 0, 2)$, $b = (3, 4, 5)$?
   b) $w_1 = (1, 2, 0)$, $w_2 = (2, 5, 0)$, $w_3 = (0, 0, 2)$, $w_4 = (0, 0, 0)$, y cualquier b?

Los problemas 19 a 37 son sobre los requerimientos para una base.

19. Si $v_1, \ldots, v_n$ son linealmente independientes, el espacio que generan tiene dimensión ____. Estos vectores son una ____ para ese espacio. Si los vectores son las columnas de cualquier matriz de m por n, entonces m es ____ que n.

20. Encuentre una base para cada uno de los siguientes subespacios de $\mathbf{R}^4$:
   a) Todos los vectores cuyas componentes son iguales.
   b) Todos los vectores tales que la suma de sus componentes es cero.
   c) Todos los vectores perpendiculares a $(1, 1, 0, 0)$ y $(1, 0, 1, 1)$.
   d) El espacio columna (en $\mathbf{R}^2$) y el espacio nulo (en $\mathbf{R}^5$) de $U = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}$.

21. Encuentre tres bases distintas para el espacio columna de la matriz U anterior. Luego encuentre dos bases distintas para el espacio renglón de U.

22. Suponga que $v_1, v_2, \ldots, v_6$ son seis vectores en $\mathbf{R}^4$.
   a) Estos vectores (generan)(no generan)(podrían generar) $\mathbf{R}^4$.
   b) Estos vectores (son)(no son)(podrían ser) linealmente independientes.
   c) Cuatro cualesquiera de esos vectores (son)(no son)(podrían ser) una base de $\mathbf{R}^4$.
   d) Si esos vectores son las columnas de A, entonces $Ax = b$ (tiene)(no tiene)(podría tener) una solución.

23. Las columnas de A son n vectores de $\mathbf{R}^m$. Si son linealmente independientes, ¿cuál es el rango de A? Si generan $\mathbf{R}^m$, ¿cuál es el rango? Si son una base de $\mathbf{R}^m$, entonces ¿qué ocurre?

24. Encuentre una base del plano $x - 2y + 3z = 0$ en $\mathbf{R}^3$. Luego encuentre una base para la intersección de ese plano con el plano xy. Luego encuentre una base de todos los vectores perpendiculares al plano.

25. Suponga que las columnas de una matriz invertible A de 5 por 5 son una base de $\mathbf{R}^5$.
   a) La ecuación $Ax = 0$ sólo tiene la solución $x = 0$ porque ____.
   b) Si b está en $\mathbf{R}^5$, entonces $Ax = b$ es resoluble porque ____.

   Conclusión: A es invertible. Su rango es 5.

26. Suponga que $\mathbf{S}$ es un subespacio pentadimensional de $\mathbf{R}^6$. ¿Falso o verdadero?
   a) Toda base de $\mathbf{S}$ puede extenderse a una base de $\mathbf{R}^6$ sumando un vector más.
   b) Toda base de $\mathbf{R}^6$ puede reducirse a una base de $\mathbf{S}$ quitando un vector.

27. U se obtuvo a partir de A restando el renglón 1 del renglón 3:

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & 1 \\ 1 & 3 & 2 \end{bmatrix} \quad \text{y} \quad U = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

   Encuentre bases de los dos espacios columna. Encuentre bases de los dos espacios renglón. Encuentre bases de los dos espacios nulos.

28. ¿Falso o verdadero? (proporcione una razón aceptable).
   a) Si las columnas de una matriz son dependientes, también lo son los renglones.
   b) El espacio columna de una matriz de 2 por 2 es el mismo que su espacio renglón.
   c) El espacio columna de una matriz de 2 por 2 tiene la misma dimensión que su espacio renglón.
   d) Las columnas de una matriz son una base para el espacio columna.

29. ¿Para cuáles números c y d el rango de las siguientes matrices es 2?

$$A = \begin{bmatrix} 1 & 2 & 5 & 0 & 5 \\ 0 & 0 & c & 2 & 2 \\ 0 & 0 & 0 & d & 2 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} c & d \\ d & c \end{bmatrix}.$$

30. Localice los pivotes, para encontrar una base del espacio columna de

$$U = \begin{bmatrix} 0 & 5 & 4 & 3 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

   Exprese cada columna que no está en la base como una combinación de las columnas de la base. También encuentre una matriz A con esta forma escalonada U, pero cuyo espacio columna sea diferente.

31. Encuentre un contraejemplo de la siguiente afirmación: Si v_1, v_2, v_3, v_4 es una base del espacio vectorial $\mathbf{R}^4$, y si $\mathbf{W}$ es un subespacio, entonces algún subconjunto de las vs es una base de $\mathbf{W}$.

32. Encuentre las dimensiones de los siguientes espacios vectoriales:
   a) El espacio de todos los vectores en $\mathbf{R}^4$ tales que la suma de sus componentes es cero.
   b) El espacio nulo de la matriz identidad de 4 por 4.
   c) El espacio de todas las matrices de 4 por 4.

33. Suponga que se sabe que la dimensión de $\mathbf{V}$ es k. Demuestre que
   a) cualesquiera k vectores independientes en $\mathbf{V}$ son una base;
   b) cualesquiera k vectores que generan $\mathbf{V}$ son una base.

   En otras palabras, si se sabe que el número de vectores es correcto, entonces una de las dos propiedades de una base implica a la segunda.

34. Demuestre que si $\mathbf{V}$ y $\mathbf{W}$ son subespacios vectoriales tridimensionales de $\mathbf{R}^5$, entonces $\mathbf{V}$ y $\mathbf{W}$ deben tener en común un vector diferente de cero. *Sugerencia*: Empiece con bases de los dos subespacios, formando seis vectores en total.

35. *¿Falso o verdadero?*
   a) Si las columnas de A son linealmente independientes, entonces $Ax = b$ tiene exactamente una solución para toda b.
   b) Una matriz de 5 por 7 nunca tiene columnas linealmente independientes.

36. Si A es una matriz de 64 por 17 con rango 11, ¿cuántos vectores independientes cumplen $Ax = 0$? ¿cuántos vectores independientes cumplen $A^{\mathrm{T}}y = 0$?

37. Encuentre una base de cada uno de los siguientes subespacios de matrices de 3 por 3:
   a) Todas las matrices diagonales.
   b) Todas las matrices simétricas ($A^{\mathrm{T}} = A$).
   c) Todas las matrices sesgadas simétricas ($A^{\mathrm{T}} = -A$).

Los problemas 38 a 42 son sobre espacios en los que los "vectores" son funciones.

38. *a*) Encuentre todas las funciones que cumplen $\frac{dy}{dx} = 0$.

b) Encuentre una función particular que cumpla $\frac{dy}{dx} = 3$.

c) Encuentre todas las funciones que cumplen $\frac{dy}{dx} = 3$.

39. El espacio coseno $\mathbf{F}_3$ contiene todas las combinaciones $y(x) = A \cos x + B \cos 2x + C \cos 3x$. Encuentre una base del subespacio que cumple $y(0) = 0$.

40. Encuentre una base para el espacio de las funciones que cumplen lo siguiente:

a) $\dfrac{dy}{dx} - 2y = 0$.

b) $\dfrac{dy}{dx} - \dfrac{y}{x} = 0$.

41. Suponga que $y_1(x), y_2(x), y_3(x)$ son tres funciones distintas de x. El espacio vectorial que pueden generar tiene dimensión 1, 2, 3. Proporcione un ejemplo de y_1, y_2, y_3 que muestre cada posibilidad.

42. Encuentre una base para el espacio de los polinomios $p(x)$ de grado ≤ 3. Encuentre una base para el subespacio de $p(1) = 0$.

43. ¡Escriba la matriz identidad de 3 por 3 como una combinación de las otras cinco matrices permutación! Luego demuestre que estas cinco matrices son linealmente independientes. (Suponga que una combinación proporciona cero, y verifique los elementos para probar que cada término es cero.) Las cinco permutaciones son una base del subespacio de matrices de 3 por 3 cuyas sumas de renglones y columnas es igual.

44. *Repaso*: ¿En cuáles de los siguientes incisos se muestran bases de $\mathbf{R}^3$?

a) $(1, 2, 0)$ y $(0, 1, -1)$.

b) $(1, 1, -1), (2, 3, 4), (4, 1, -1), (0, 1, -1)$.

c) $(1, 2, 2), (-1, 2, 1), (0, 8, 0)$.

d) $(1, 2, 2), (-1, 2, 1), (0, 8, 6)$.

45. *Repaso*: Suponga que A es de 5 por 4 con rango 4. Demuestre que $Ax = b$ no tiene solución cuando la matriz $[A\ b]$ de 5 por 5 es invertible. Demuestre que $Ax = b$ es resoluble cuando $[A\ b]$ es singular.

2.4 LOS CUATRO SUBESPACIOS FUNDAMENTALES

La sección anterior se centró más en definiciones que en construcciones. Se sabe lo que es una base, pero no cómo encontrar una. Ahora, empezando con una descripción explícita de un subespacio, sería conveniente calcular una base explícita.

Los subespacios pueden describirse en dos formas. Primero, puede proporcionarse un subconjunto de vectores que generan el espacio. (*Ejemplo*: Las columnas generan el espacio columna.) Segundo, pueden especificarse las condiciones que deben cumplir los vectores en el espacio. (*Ejemplo*: El espacio nulo consta de todos los vectores que cumplen $Ax = 0$.)

La primera descripción puede incluir vectores inútiles (columnas dependientes). La segunda descripción puede incluir condiciones repetidas (renglones dependientes). No es posible escribir una base por inspección, de modo que se requiere un procedimiento sistemático.

El lector puede adivinar cuál debe ser el procedimiento. Cuando la eliminación sobre A produce una matriz escalonada U o una reducida R, para cada uno de los subespacios asociados con A se encuentra una base. Luego tiene que considerarse el caso extremo de **rango total**:

Cuando el rango es lo más grande posible, $\boldsymbol{r = n}$ *o* $\boldsymbol{r = m}$ *o* $\boldsymbol{r = m = n}$, *la matriz tiene una* **inversa izquierda** B, *o una* **inversa derecha** C *o una* A^{-1} *por* **ambos lados.**

Con la finalidad de organizar todo el análisis, cada uno de los cuatro subespacios se abordará a su vez. Dos de ellos son conocidos y dos son nuevos.

1. El ***espacio columna*** de A se denota por $C(A)$. Su dimensión es el rango r.
2. El ***espacio nulo*** de A se denota por $N(A)$. Su dimensión es $n - r$.
3. El ***espacio renglón*** de A es el *espacio columna* de A^{T}. Es $C(A^{\mathrm{T}})$, y es generado por los renglones de A. Su dimensión también es r.
4. El ***espacio nulo izquierdo*** de A es el *espacio nulo* de A^{T}. Contiene a todos los vectores y, tales que $A^{\mathrm{T}}y = 0$, y se escribe $N(A^{\mathrm{T}})$. Su dimensión es ____.

La cuestión sobre los dos últimos subespacios es que *provienen de* A^{T}. Si A es una matriz de m por n, es posible ver cuáles espacios "huéspedes" contienen a los cuatro subespacios, al observar el número de componentes:

El espacio nulo $N(A)$ y el espacio renglón $C(A^{\mathrm{T}})$ son subespacios de $\mathbf{R}^n$.
El espacio nulo izquierdo $N(A^{\mathrm{T}})$ y el espacio columna $C(A)$ son subespacios de $\mathbf{R}^m$.

Los renglones tienen n componentes y las columnas tienen m. Para una matriz sencilla como

$$A = U = R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

el espacio columna es la recta que pasa por $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$. El espacio renglón es la recta que pasa por $[1\ 0\ 0]^{\mathrm{T}}$. Está en $\mathbf{R}^3$. El espacio nulo es un plano en $\mathbf{R}^3$ y el espacio nulo izquierdo es una recta en $\mathbf{R}^2$:

$$N(A) \text{ contiene } \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ y } \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad N(A^{\mathrm{T}}) \text{ contiene } \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Observe que todos los vectores son vectores columna. Incluso los renglones están traspuestos, y el espacio renglón de A es el espacio *columna* de A^{T}. Nuestro problema es relacionar los cuatro espacios para U (después de la eliminación) con los cuatro espacios para A:

Ejemplo básico

$$U = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{proviene de} \quad A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix}.$$

Como novedad, los cuatro espacios se considerarán en un orden más interesante.

3. El espacio renglón de A Para cada matriz escalonada como U, el espacio renglón es evidente. Contiene a todas las combinaciones de los renglones, como ocurre con cualquier espacio renglón, aunque aquí el tercer renglón no contribuye a nada. Los dos primeros renglones son una base del espacio renglón. Una regla semejante es válida para toda matriz escalonada U o R, con r pivotes y r renglones diferentes de cero: ***los renglones diferentes de cero son una base, y la dimensión del espacio renglón es r***. Esto facilita tratar con la matriz original A.

2M El espacio renglón de A tiene la misma dimensión r que el espacio renglón de U, y tiene las mismas bases, porque ***los espacios renglón de A y U (y R) son iguales.***

La razón es que cada operación elemental deja sin cambio el espacio renglón. Los renglones en U son combinaciones de los renglones originales en A. En consecuencia, el espacio renglón de U no contiene nada nuevo. Al mismo tiempo, debido a que cada paso puede invertirse, nada se pierde; los renglones de A pueden recuperarse de U. Es cierto que los renglones de A y U son distintos, pero las *combinaciones* de los renglones son idénticas: ¡*mismo espacio*!

Observe que no se empezó con los m renglones de A, que generan el espacio renglón, y que se eliminan $m - r$ de ellos para terminar con una base. Según 2L, hubiera podido hacerse lo anterior. Sin embargo, podría ser difícil decidir cuáles renglones preservar, y cuáles eliminar, por lo que fue más fácil considerar los renglones de U diferentes de cero.

2. El espacio nulo de A La eliminación simplifica un sistema de ecuaciones lineales sin cambiar las soluciones. El sistema $Ax = 0$ se reduce al sistema $Ux = 0$, y este proceso es reversible. ***El espacio nulo de A es el mismo que el espacio nulo de U y R.*** Sólo r de las ecuaciones $Ax = b$ son independientes. Al elegir las $n - r$ "soluciones especiales" de $Ax = 0$ se obtiene una base definida para el espacio nulo:

2N La dimensión del espacio nulo $N(A)$ es $n - r$. Las "soluciones especiales" son una base: a cada variable libre se asigna el valor 1, mientras las otras variables libres son 0. Así, $Ax = 0$ o $Ux = 0$ o $Rx = 0$ proporcionan las variables pivote por sustitución hacia atrás.

Ésta es exactamente la forma en que se ha venido resolviendo $Ux = 0$. El ejemplo básico anterior tiene pivotes en las columnas 1 y 3. En consecuencia, sus variables libres son las columnas segunda y cuarta, v y y. La base del espacio nulo es

$$\textbf{Soluciones especiales} \qquad \begin{matrix} v = 1 \\ y = 0 \end{matrix} \quad x_1 = \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}; \qquad \begin{matrix} v = 0 \\ y = 1 \end{matrix} \quad x_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}.$$

Cualquier combinación $c_1x_1 + c_2x_2$ tiene a c_1 en su componente v, y a c_2 en su componente y. La única forma en que $c_1x_1 + c_2x_2 = 0$ es que $c_1 = c_2 = 0$, de modo que estos vectores son independientes. También generan el espacio nulo; la solución completa es $vx_1 + yx_2$. Así, los $n - r = 4 - 2$ vectores son una base.

El espacio nulo también se denomina *kernel* de A, y su dimensión $n - r$ es la *nulidad*.

1. El espacio columna de A El espacio columna a veces se denomina **rango**. Esto es consistente con la idea de costumbre que se tiene respecto del contradominio como el conjunto de todos los valores posibles $f(x)$; x está en el dominio y $f(x)$ está en el contradominio. En nuestro caso, la función es $f(x) = Ax$. Su dominio consta de todas las x en $\mathbf{R}^n$; su contradominio consta de todos los vectores posibles Ax, que es el espacio columna. (En una edición previa a ésta se denominó $R(A)$.)

El problema consiste en encontrar base para los espacios columna de U y A. ***Estos espacios son distintos*** (¡simplemente observe las matrices!) aunque su dimensión es la misma.

Las columnas primera y tercera de U son una base de su espacio columna. Se trata de ***las columnas que contienen a los pivotes***. Cualquier otra columna es una combinación de estas dos. Además, lo mismo es cierto para la A original; aun cuando sus columnas son diferentes. ***Las columnas pivote de A son una base de su espacio columna.*** La segunda columna es tres veces la primera, justo como en U. La cuarta columna es igual a (columna 3) − (columna 1). El mismo espacio nulo indica estas dependencias.

Lo anterior se debe a que $Ax = 0$ *exactamente cuando* $Ux = 0$. Los dos sistemas son equivalentes y tienen las mismas soluciones. La cuarta columna de U también era (columna 3) − (columna 1). Toda dependencia lineal $Ax = 0$ entre las columnas de A es compensada por una dependencia $Ux = 0$ entre las columnas de U, con exactamente los mismos coeficientes. *Si un conjunto de columnas de A es independiente, entonces también lo son las columnas correspondientes de U, y viceversa.*

Con la finalidad de encontrar una base del espacio columna $C(A)$, se usa lo que ya se ha hecho para U. Las r columnas que contienen a los pivotes son una base del espacio columna de U. Estas mismas r columnas se elegirán en A:

2O La dimensión del espacio columna $C(A)$ es igual al rango r, que también es igual a la dimensión del espacio renglón: ***El número de columnas independientes es igual al número de renglones independientes.*** Una base para $C(A)$ se forma con las r columnas en A que corresponden, en U, a las columnas que contienen a los pivotes.

¡El espacio renglón y el espacio columna tienen la misma dimensión r! Este es uno de los teoremas más importantes en álgebra lineal. A menudo se abrevia como "***rango de los renglones = rango de las columnas***". Expresa un resultado que, para una matriz aleatoria de 10 por 12 no es para nada evidente. También dice algo sobre las matrices cuadradas: *Si los renglones de una matriz cuadrada son linealmente independientes, entonces también lo son las columnas* (y viceversa). De nuevo, esto no parece autoevidente (por lo menos, no para el autor).

Para ver otra vez que la dimensión del espacio columna y del espacio renglón de U es r, considere una situación típica con rango $r = 3$. La matriz escalonada U ciertamente tiene tres renglones independientes:

$$U = \begin{bmatrix} d_1 & * & * & * & * & * \\ 0 & 0 & 0 & d_2 & * & * \\ 0 & 0 & 0 & 0 & 0 & d_3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Se afirma que U también tiene tres columnas independientes y no más que tres. Estas columnas sólo tienen tres componentes diferentes de cero. Si puede demostrarse que las columnas pivote: la primera, la cuarta, y la sexta, son linealmente independientes, entonces debe haber una base (¡para el espacio columna de U, no el de A!) Suponga que una combinación de estas columnas pivote produjo cero:

$$c_1 \begin{bmatrix} d_1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} * \\ d_2 \\ 0 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} * \\ * \\ d_3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Trabajando hacia arriba como de costumbre, c_3 debe ser cero porque el pivote $d_3 \neq 0$; así, c_2 debe ser cero porque el pivote $d_2 \neq 0$; y finalmente $c_1 = 0$. Con esto se establece la independencia y se completa la demostración. Como $Ax = 0$ si y sólo si $Ux = 0$, las columnas primera, cuarta y sexta de A, sin importar cuál era la matriz original A, que ni siquiera se conoció en este ejemplo, son una base de $C(A)$.

El espacio columna y el espacio renglón se vuelven evidentes después de la eliminación sobre A. A continuación se abordará el cuarto subespacio fundamental, que se ha mantenido silenciosamente fuera de la vista. Debido a que los tres primeros espacios eran $C(A)$, $N(A)$ y $C(A^T)$, el cuarto espacio debe ser $N(A^T)$. Es el espacio nulo de la traspuesta, o el ***espacio nulo izquierdo*** de A. $A^Ty = 0$ significa que $y^TA = 0$, y el vector aparece en el miembro izquierdo de A.

4. El espacio nulo izquierdo de A (= espacio nulo de A^T) Si A es una matriz de m por n, entonces A^T es de n por m. Su espacio nulo es un subespacio de $\mathbf{R}^m$; el vector y tiene m componentes. Si se escribe como $y^TA = 0$, estas componentes multiplican los *renglones* de A para producir el renglón cero:

$$y^TA = \begin{bmatrix} y_1 & \cdots & y_m \end{bmatrix} \begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} = \begin{bmatrix} 0 & \cdots & 0 \end{bmatrix}.$$

La dimensión de este espacio nulo $N(A^T)$ es fácil de encontrar. Para *cualquier* matriz, ***el número de variables pivote más el número de variables libres debe corresponder al número total de columnas***. Para A, era $r + (n - r) = n$. En otras palabras, el rango más las nulidades igual a n:

$$\textit{dimensión de } C(A) + \textit{dimensión de } N(A) = \textit{número de columnas.}$$

Esta ley se aplica igualmente a A^T, que tiene m columnas. A^T es tan buena matriz como A. Pero la dimensión de su espacio columna también es r, de modo que

$$r + \text{dimensión}\,(N(A^T)) = m. \qquad (1)$$

2P La dimensión del espacio nulo izquierdo $N(A^T)$ es $m - r$.

Las $m - r$ soluciones de $y^TA = 0$ se esconden en alguna parte durante la eliminación. Los renglones de A se combinan para producir los $m - r$ *renglones cero* de U. Se empieza con $PA = LU$, o con $L^{-1}PA = U$. Los últimos $m - r$ renglones de la matriz invertible $L^{-1}P$ deben ser una base de las ys en el espacio nulo izquierdo, ya que multiplican a A para producir los renglones cero en U.

En el ejemplo de 3 por 4, el renglón cero era el renglón 3 − 2(renglón 2) + 5(renglón 1). En consecuencia, las componentes de y son 5, −2, 1. Esta es la misma combinación que en $b_3 - 2b_2 + 5b_1$ en el miembro derecho, lo cual lleva a $0 = 0$ como la ecuación final. Ese vector y es una base para el espacio nulo izquierdo, cuya dimensión es $m - r = 3 - 2 = 1$. Es el último renglón de $L^{-1}P$, y produce el renglón cero en U, y a menudo puede verse sin necesidad de calcular L^{-1}. Cuando se está desesperado, siempre es posible resolver precisamente $A^Ty = 0$.

Me doy cuenta de que hasta el momento, en este libro, no se ha proporcionado ninguna razón para tener cuidado sobre $N(A^T)$. Es correcto pero no convincente si en cursivas se escribe que *el espacio nulo izquierdo siempre es importante*. En la siguiente sección se mejora el tema de encontrar un significado físico para y, a partir de la ley de la corriente de Kirchhoff.

Ahora ya se conocen las dimensiones de los cuatro espacios. Pueden resumirse en una tabla, y aún así falta mucho para poder indicarlas como el

Teorema fundamental del álgebra lineal, parte I

1. $C(A)$ = espacio columna de A; dimensión r.
2. $N(A)$ = espacio nulo de A; dimensión $n - r$.

3. $C(A^T)$ = espacio renglón de A; dimensión r.
4. $N(A^T)$ = espacio nulo izquierdo de A; dimensión $m - r$.

Ejemplo 1 $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ tiene $m = n = 2$, y rango $r = 1$.

1. El ***espacio columna*** contiene a todos los múltiplos de $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$. La segunda columna está en la misma dirección y no contribuye con nada nuevo.
2. El ***espacio nulo*** contiene a todos los múltiplos de $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$. Este vector satisface $Ax = 0$.
3. El ***espacio renglón*** contiene a todos los múltiplos de $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Lo escribo como un vector columna porque, hablando estrictamente, está en el espacio columna de A^T.
4. El ***espacio nulo izquierdo*** contiene a todos los múltiplos de $y = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$. La suma de los renglones de A con coeficientes −3 y 1 es cero, de modo que $A^Ty = 0$.

En este ejemplo *todos los cuatro subespacios son rectas*. Esto es accidental, y proviene de $r = 1$ y $n - r = 1$ y $m - r = 1$. En la figura 2.5 se muestra que dos pares de rectas son perpendiculares. ¡Esto no es accidental!

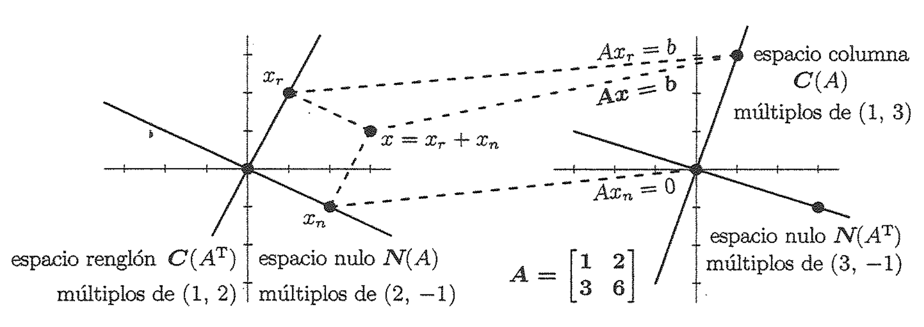


Figura 2.5 Los cuatro subespacios fundamentales (rectas) de la matriz singular A.

Si el último elemento de A se cambia de 6 a 7, todas las dimensiones son diferentes. La dimensión del espacio columna y el espacio renglón es $r = 2$. El espacio nulo y el espacio nulo izquierdo sólo contienen a los vectores $x = 0$ y $y = 0$. *La matriz es invertible.*

Existencia de inversas

Se sabe que si A tiene una inversa izquierda ($BA = I$) y una inversa derecha ($AC = I$), entonces las dos inversas son iguales: $B = B(AC) = (BA)C = C$. Así, por el rango de una matriz, es fácil decidir cuáles matrices tienen en realidad estas inversas. En términos generales, ***una inversa existe sólo cuando el rango es lo más grande posible***.

El rango siempre cumple $r \le m$ y también $r \le n$. Una matriz de m por n no puede tener más de m renglones independientes o n columnas independientes. No hay espacio para más de m pivotes, o más de n. Se quiere demostrar que cuando $r = m$ hay una inversa derecha, y que $Ax = b$ siempre tiene una solución única. Cuando $r = n$, hay una inversa izquierda y la solución (*si existe*) es única.

Sólo una matriz cuadrada puede tener ambos $r = m$ y $r = n$, y en consecuencia sólo una matriz cuadrada puede tener existencia y unicidad. Sólo una matriz cuadrada tiene una inversa por ambos lados.

2Q EXISTENCIA: Rango renglón total $r = m$. $Ax = b$ tiene *por lo menos* una solución x para toda b si y sólo si las columnas generan $\mathbf{R}^m$. Así, A tiene una **inversa derecha** C tal que $AC = I_m$ (de m por m). Esto sólo es posible si $m \leq n$.

UNICIDAD: Rango columna total $r = n$. $Ax = b$ tiene *cuando mucho* una solución x para toda b si y sólo si las columnas son linealmente independientes. Entonces A tiene una **inversa izquierda** B de n por m tal que $BA = I_n$. Esto sólo es posible si $m \geq n$.

En el caso de existencia, una solución posible es $x = Cb$, ya que entonces $Ax = ACb = b$. Sin embargo, hay otras soluciones si hay otras inversas derechas. El número de soluciones cuando las columnas generan $\mathbf{R}^m$ es 1 o ∞.

En el caso de unicidad, si hay una solución de $Ax = b$, debe ser $x = BAx = Bb$. Pero puede no haber solución. El número de soluciones es 0 o 1.

Hay fórmulas sencillas para encontrar las mejores inversas izquierda y derecha, en caso de existir:

$$\textbf{Inversas por un lado} \qquad B = (A^TA)^{-1}A^T \text{ y } C = A^T(AA^T)^{-1}.$$

Ciertamente $BA = I$ y $AC = I$. Lo que no es tan cierto es que A^TA y AA^T son realmente invertibles. En el capítulo 3 se demostrará que A^TA tiene inversa si el rango es n, y AA^T tiene inversa cuando el rango es m. Así, las fórmulas tienen sentido exactamente cuando el rango es lo más grande posible, y se encuentran las inversas por un lado.

Ejemplo 2 Considere una matriz sencilla de 2 por 3 con rango 2:

$$A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix}.$$

Debido a que $r = m = 2$, el teorema garantiza una inversa derecha C:

$$AC = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{5} \\ c_{31} & c_{32} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Hay muchas inversas derechas porque el último renglón de C es completamente arbitrario. Este es un caso de existencia pero no de unicidad. La matriz A no tiene inversa izquierda porque la última columna de BA es ciertamente cero. La inversa derecha específica $C = A^T(AA^T)^{-1}$ elige que c_{31} y c_{32} sean cero:

$$\textbf{La mejor inversa derecha} \qquad A^T(AA^T)^{-1} = \begin{bmatrix} 4 & 0 \\ 0 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{16} & 0 \\ 0 & \frac{1}{25} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{5} \\ 0 & 0 \end{bmatrix} = C.$$

Esta es la *seudoinversa*: una forma de escoger la mejor C en la sección 6.3. La traspuesta de A, lleva a un ejemplo de una infinidad de inversas *izquierdas*:

$$BA^T = \begin{bmatrix} \frac{1}{4} & 0 & b_{13} \\ 0 & \frac{1}{5} & b_{23} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 5 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Ahora, la que es completamente arbitraria es la última columna de B. La mejor inversa izquierda (también la seudoinversa) tiene $b_{13} = b_{23} = 0$. Este es un "caso de unicidad", cuando el rango es $r = n$.

No hay variables libres, ya que $n - r = 0$. Si existe una solución, es única. El lector puede ver cuándo este ejemplo tiene una solución o no tiene solución.

$$\begin{bmatrix} 4 & 0 \\ 0 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \text{ es resoluble exactamente cuando } b_3 = 0.$$

Una matriz rectangular no puede tener tanto existencia como unicidad. Si m es diferente de n, entonces no puede cumplirse $r = m$ y $r = n$.

Una matriz cuadrada constituye el caso opuesto. Si $m = n$, no puede cumplirse una propiedad *sin* la otra. Una matriz cuadrada tiene una inversa izquierda si y sólo si tiene una inversa derecha. Sólo hay una inversa; a saber, $B = C = A^{-1}$. *Cuando la matriz es cuadrada, la existencia implica la unicidad y la unicidad implica la existencia.* La condición para invertibilidad es el **rango total**: $r = m = n$. Cada una de las siguientes condiciones es una prueba necesaria y suficiente:

1. Las columnas generan $\mathbf{R}^n$, de modo que $Ax = b$ tiene por lo menos una solución para toda b.
2. Las columnas son independientes, de modo que $Ax = 0$ sólo tiene la solución $x = 0$.

Esta lista puede hacerse mucho más grande, especialmente si se piensa en los capítulos ulteriores. Cada condición es equivalente a cualquier otra, lo cual asegura que A es invertible.

3. Los renglones de A generan $\mathbf{R}^n$.
4. Los renglones son linealmente independientes.
5. Es posible completar la eliminación: $PA = LDU$, con todos los n pivotes.
6. El determinante de A es diferente de cero.
7. El cero no es un valor característico de A.
8. A^TA es positiva definida.

A continuación se presenta una aplicación típica a polinomios $P(t)$ de grado $n - 1$. El único de estos polinomios que se hace cero en $t_1, \ldots, t_n$ es $P(t) \equiv 0$. Ningún otro polinomio de grado $n - 1$ puede tener n raíces. Esto es unicidad, e implica existencia: Dados valores cualesquiera $b_1, \ldots, b_n$, *existe* un polinomio de grado $n - 1$ que interpola estos valores: $P(t_i) = b_i$. La cuestión es que está tratando con una matriz cuadrada; el número n de coeficientes en $P(t) = x_1 + x_2t + \cdots + x_nt^{n-1}$ corresponde al número de ecuaciones:

$$\begin{matrix}\textbf{Interpolación} \\ \boldsymbol{P(t_i) = b_i}\end{matrix} \qquad \begin{bmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^{n-1} \\ 1 & t_2 & t_2^2 & \cdots & t_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & t_n & t_n^2 & \cdots & t_n^{n-1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Esta *matriz de Vandermonde* es de n por n y con rango total. $Ax = b$ siempre tiene una solución: un polinomio puede pasar por cualquier b_i en puntos distintos t_i. Después se encontrará realmente el determinante de A; no es cero.

Matrices de rango 1

Por último llega el caso más sencillo, cuando el rango es lo más *pequeño* posible (excepto por la matriz cero con rango 0). Uno de los temas básicos de las matemáticas es, dado algo

complicado, demostrar cómo puede descomponerse en piezas sencillas. Para el álgebra lineal, las piezas sencillas son las matrices de ***rango* 1**.

$$\textbf{Rango 1} \quad A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{bmatrix} \text{ tiene } r = 1.$$

Cada renglón es un múltiplo del primer renglón, de modo que el espacio renglón es unidimensional. De hecho es posible escribir toda la matriz como *el producto de un vector columna y un vector renglón*:

$$A = \textbf{(columna)(renglón)} \quad \begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \\ -1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \end{bmatrix}.$$

El producto de una matriz de 4 por 1 y una matriz de 1 por 3 es una matriz de 4 por 3. *El rango de este producto es* 1. Al mismo tiempo, las columnas son todos los múltiplos del mismo vector columna; el espacio columna comparte la dimensión $r = 1$, y se reduce a una recta.

***Toda matriz con rango* 1 *tiene la forma sencilla* $A = uv^T$ = *columna por renglón*.**

Todos los renglones son múltiplos del mismo vector v^T, y todas las columnas son múltiplos de u. El espacio renglón y el espacio columna son rectas: el caso más sencillo.

Conjunto de problemas 2.4

1. Falso o verdadero: Si $m = n$, entonces el espacio renglón de A es igual al espacio columna. Si $m < n$, entonces el espacio nulo tiene una dimensión mayor que ____.

2. Encuentre la dimensión, y construya una base para los cuatro subespacios asociados con cada una de las siguientes matrices

$$A = \begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 2 & 8 & 0 \end{bmatrix} \quad \text{y} \quad U = \begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

3. Encuentre la dimensión y una base para los cuatro subespacios fundamentales de

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix} \quad \text{y} \quad U = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

4. Describa los cuatro subespacios en el espacio tridimensional asociados con

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

5. Si el producto AB es la matriz cero, $AB = 0$, demuestre que el espacio columna asociado de B está contenido en el espacio nulo de A. (También, el espacio renglón de A está en el espacio nulo izquierdo de B, ya que cada renglón de A multiplica B para obtener un renglón cero.)

6. Suponga que A es cualquier matriz de m por n de rango r. ¿En qué condiciones sobre estos números

a) A tiene una inversa por ambos lados: $AA^{-1} = A^{-1}A = I$?
b) $Ax = b$ tiene una *infinidad* de soluciones *para toda* b?

7. ¿Por qué no existe ninguna matriz cero cuyos espacios renglón y nulo contengan a (1, 1, 1)?

8. Suponga que la única solución de $Ax = 0$ (m ecuaciones en n incógnitas) es $x = 0$. ¿Cuál es el rango y por qué? Las columnas de A son linealmente ____.

9. Encuentre una matriz de 1 por 3 cuyo espacio nulo conste de todos los vectores en $\mathbf{R}^3$ tales que $x_1 + 2x_2 + 4x_3 = 0$. Encuentre una matriz de 3 por 3 con el mismo espacio nulo.

10. Si $Ax = b$ siempre tiene por lo menos una solución, demuestre que la única solución de $A^Ty = 0$ es $y = 0$. *Sugerencia*: ¿Cuál es el rango?

11. Si $Ax = 0$ tiene una solución diferente de cero, demuestre que $A^Ty = f$ falla en ser resoluble para algunos miembros derechos f. Construya un ejemplo de A y f.

12. Encuentre el rango de A, y escriba la matriz como $A = uv^T$:

$$A = \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 6 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 2 & -2 \\ 6 & -6 \end{bmatrix}.$$

13. Si se proporcionan a, b, c con $a \neq 0$, escoja d de modo que el rango de

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = uv^T$$

sea 1. ¿Cuáles son los pivotes?

14. Encuentre una inversa izquierda y/o una inversa derecha (cuando existan) para

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{y} \quad M = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad T = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}.$$

15. Si las columnas de A son linealmente independientes (A es de m por n), entonces el rango es ____, el espacio nulo es ____, y existe una inversa ____.

16. (*Una paradoja*) Suponga que A tiene una inversa derecha B. Así, $AB = I$ conduce a $A^TAB = A^T$ o $B = (A^TA)^{-1}A^T$. Sin embargo, satisface $BA = I$; es una inversa *izquierda*. ¿Cuál paso no está justificado?

17. Encuentre una matriz A cuyo espacio renglón sea $\mathbf{V}$, y una matriz B cuyo espacio nulo sea $\mathbf{V}$, si $\mathbf{V}$ es el subespacio generado por

$$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix}.$$

18. Encuentre una base de cada uno de los cuatro subespacios de

$$A = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 4 & 6 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

19. Si A tiene los mismos cuatro subespacios fundamentales que B, ¿es cierto que $A = cB$?

20. *a*) Si el rango de una matriz de 7 por 9 es 5, ¿cuáles son las dimensiones de los cuatro subespacios? ¿Cuánto es la suma de los cuatro subespacios?
b) Si el rango de una matriz de 3 por 4 es 3, ¿cuáles son el espacio columna y el espacio nulo izquierdo?

21. Construya una matriz con la propiedad requerida, o explique por qué no es posible.
a) El espacio columna contiene a $\begin{bmatrix}1\\1\\0\end{bmatrix}$, $\begin{bmatrix}0\\0\\1\end{bmatrix}$, el espacio renglón contiene a $\begin{bmatrix}1\\2\end{bmatrix}$, $\begin{bmatrix}2\\5\end{bmatrix}$.
b) El espacio columna tiene como base a $\begin{bmatrix}1\\2\\3\end{bmatrix}$, el espacio nulo tiene como base a $\begin{bmatrix}3\\2\\1\end{bmatrix}$.
c) La dimensión del espacio nulo $= 1 +$ dimensión del espacio nulo izquierdo.
d) El espacio nulo izquierdo contiene a $\begin{bmatrix}1\\3\end{bmatrix}$ el espacio renglón contiene a $\begin{bmatrix}3\\1\end{bmatrix}$.
e) Espacio renglón = espacio columna, espacio nulo $\neq$ espacio nulo izquierdo.

22. Sin eliminación, encuentre las dimensiones y bases de los cuatro subespacios de
$$A = \begin{bmatrix}0&3&3&3\\0&0&0&0\\0&1&0&1\end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix}1&1\\4&4\\5&5\end{bmatrix}.$$

23. Suponga que una matriz A de 3 por 3 es invertible. Escriba bases de los cuatro subespacios de A, y también para la matriz de 3 por 6 $B = [A\ A]$.

24. ¿Cuáles son las dimensiones de los cuatro subespacios de A, B, y C, si I es la matriz identidad de 3 por 3, y 0 es la matriz cero de 3 por 2?
$$A = \begin{bmatrix}I & 0\end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix}I & I\\ 0^{\mathrm{T}} & 0^{\mathrm{T}}\end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix}0\end{bmatrix}.$$

25. ¿Cuáles subespacios son iguales para las siguientes matrices de tamaños distintos?
$$a)\ [A] \quad \text{y} \quad \begin{bmatrix}A\\A\end{bmatrix}. \qquad b)\ \begin{bmatrix}A\\A\end{bmatrix} \quad \text{y} \quad \begin{bmatrix}A&A\\A&A\end{bmatrix}.$$
Demuestre que el rango de estas tres matrices es el mismo: r.

26. Si los elementos de una matriz de 3 por 3 se escogen aleatoriamente entre 0 y 1, ¿cuáles son las dimensiones más probables de los cuatro subespacios? ¿Qué ocurre si la matriz es de 3 por 5?

27. (Importante) A es una matriz de m por n con rango r. Suponga que hay miembros derechos b para los cuales $Ax = b$ *no tiene solución*.
a) ¿Cuáles desigualdades ($<$ o $\leq$) deben ser ciertas entre m, n, y r?
b) ¿Cómo se sabe que $A^{\mathrm{T}}y = 0$ tiene una solución diferente de cero?

28. Construya una matriz con (1, 0, 1) y (1, 2, 0) como una base para su espacio renglón y su espacio columna. ¿Por qué esta base no puede ser una base del espacio renglón y del espacio nulo?

29. Sin calcular A, encuentre bases de los cuatro subespacios fundamentales:
$$A = \begin{bmatrix}1&0&0\\6&1&0\\9&8&1\end{bmatrix}\begin{bmatrix}1&2&3&4\\0&1&2&3\\0&0&1&2\end{bmatrix}.$$

30. Si se intercambian los dos primeros renglones de una matriz A, ¿cuáles de los cuatro subespacios permanecen igual? Si $y = (1, 2, 3, 4)$ está en el espacio nulo izquierdo de A, escriba un vector en el espacio nulo izquierdo de la nueva matriz.

31. Explique por qué $v = (1, 0, -1)$ no puede ser un renglón de A y también estar en el espacio nulo.

32. Describa los cuatro subespacios de $\mathbf{R}^3$ asociados con
$$A = \begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix} \quad \text{y} \quad I + A = \begin{bmatrix}1&1&0\\0&1&1\\0&0&1\end{bmatrix}.$$

33. (Espacio nulo izquierdo) Sume la columna extra b, y reduzca A a forma escalonada:
$$\begin{bmatrix}A & b\end{bmatrix} = \begin{bmatrix}1&2&3&b_1\\4&5&6&b_2\\7&8&9&b_3\end{bmatrix} \to \begin{bmatrix}1&2&3&b_1\\0&-3&-6&b_2-4b_1\\0&0&0&b_3-2b_2+b_1\end{bmatrix}.$$
Una combinación de los renglones de A produjo el renglón cero. ¿Qué combinación es? (Observe $b_3 - 2b_2 + b_1$ en el miembro derecho.) ¿Cuáles vectores están en el espacio nulo de A^{T}, y cuáles están en el espacio nulo de A?

34. Siga el método del problema 33 para reducir A a forma escalonada, y busque los renglones cero. La columna b indica cuáles combinaciones de los renglones se tomaron.
$$a)\ \begin{bmatrix}1&2&b_1\\3&4&b_2\\4&6&b_3\end{bmatrix}. \qquad b)\ \begin{bmatrix}1&2&b_1\\2&3&b_2\\2&4&b_3\\2&5&b_4\end{bmatrix}.$$
A partir de la columna b después de la eliminación, lea $m - r$ vectores de la base en el espacio nulo izquierdo de A (combinaciones de renglones que proporcionan cero).

35. Suponga que A es la suma de dos matrices de rango 1: $A = uv^{\mathrm{T}} + wz^{\mathrm{T}}$.
a) ¿Cuáles vectores generan el espacio columna de A?
b) ¿Cuáles vectores generan el espacio renglón de A?
c) El rango es menos de 2 si ____ o si ____.
d) Calcule A y su rango, si $u = z = (1, 0, 0)$ y $v = w = (0, 0, 1)$.

36. Sin multiplicar las matrices, encuentre bases de los espacios renglón y columna de A:
$$A = \begin{bmatrix}1&2\\4&5\\2&7\end{bmatrix}\begin{bmatrix}3&0&3\\1&1&2\end{bmatrix}.$$
¿Cómo se sabe a partir de estas formas que A no es invertible?

37. ¿Falso o verdadero? (Según corresponda, proporcione una razón o un contraejemplo.)
a) A y A^{T} tienen el mismo número de pivotes.
b) A y A^{T} tienen el mismo espacio nulo izquierdo.
c) Si el espacio renglón es igual al espacio columna, entonces $A^{\mathrm{T}} = A$.
d) Si $A^{\mathrm{T}} = -A$, entonces el espacio renglón de A es igual al espacio columna.

38. Si $AB = 0$, las columnas de B están en el espacio nulo de A. Si estos vectores están en $\mathbf{R}^n$, demuestre que rango(A) + rango(B) $\leq n$.

39. ¿Es posible completar un juego de "gato" (tic-tac-toe) (5 unos y 4 ceros en A) de modo que rango(A) = 2, pero que en ningún lado se tenga una jugada ganadora?

40. Construya una matriz de 2 por 3 de rango 1. Copie la figura 2.5, y ponga un vector en cada subespacio (dos en el espacio nulo). ¿Cuáles vectores son ortogonales?

41. Vuelva a dibujar la figura 2.5 para una matriz de 3 por 2 de rango $r = 2$. ¿Cuál subespacio es $\mathbf{Z}$ (sólo el vector cero)? La parte del espacio nulo de cualquier vector x en $\mathbf{R}^2$ es $x_n =$ ____.

2.5 GRÁFICAS Y REDES

No estoy muy contento con la matriz de 3 por 4 en la sección previa. Desde un punto de vista teórico, era bastante satisfactoria; los cuatro subespacios eran calculables y sus dimensiones r, $n - r$, r, $m - r$ eran diferentes de cero. Sin embargo, el ejemplo no fue producto de una verdadera aplicación. No mostró cuán fundamentales realmente son estos subespacios.

En esta sección se presenta una clase de matrices rectangulares con dos ventajas. Son sencillas y son importantes. Son ***matrices de incidencia de gráficas***, y cada elemento es 1, -1, o 0. Lo extraordinario es que lo mismo es cierto de L, U y vectores de la base para los cuatro subespacios. Estos subespacios desempeñan un papel central en la teoría de redes. Se recalca que el término "gráfica" no se refiere a la gráfica de una función (como una parábola para $y = x^2$). Hay un segundo significado, completamente distinto, más próximo a las ciencias de la computación que al cálculo, que es fácil explicar. *Esta sección es opcional*, pero constituye una oportunidad para ver en acción a las matrices rectangulares, así como la forma en que la matriz simétrica cuadrada A^TA aparece al final.

Una ***gráfica*** consta de un conjunto de vértices o *nodos* y un conjunto de *aristas* que unen a los nodos. La gráfica de la figura 2.6 tiene 4 nodos y 5 aristas. No tiene una arista entre los nodos 1 y 4 (y las aristas de un nodo hacia sí mismo están prohibidas). Esta gráfica es *dirigida*, debido a la flecha que hay en cada arista.

La ***matriz de incidencia arista-nodo*** es de 5 por 4, con un renglón para cada arista. ***Si la arista va del nodo j al nodo k, entonces ese renglón tiene −1 en la columna j y +1 en la columna k.*** La matriz de incidencia A se muestra junto a la gráfica (aunque también podría recuperarse la gráfica si sólo se contara con A). El renglón 1 muestra la arista del nodo 1 al nodo 2. El renglón 5 proviene de la quinta arista, que va del nodo 3 al nodo 4.

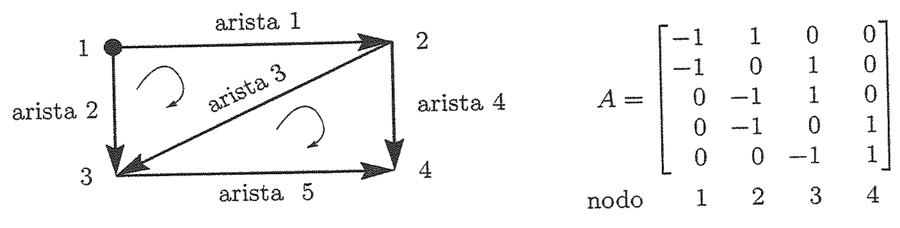


Figura 2.6 Gráfica dirigida (5 aristas, 4 nodos, 2 circuitos) y su matriz de incidencia A.

Observe las columnas de A. La columna 3 proporciona información sobre el nodo 3: indica qué arista entra y qué arista sale. Las aristas 2 y 3 entran, la arista 5 sale (con el signo menos). Algunas veces A se denomina matriz *conectividad* o matriz *topología*. Cuando la gráfica tiene m aristas y n nodos, A es de m por n (y normalmente $m > n$). Su traspuesta es la matriz de incidencia "nodo-arista".

Cada uno de los cuatro subespacios fundamentales tiene un significado en términos de la gráfica. Es posible hacer algo de álgebra lineal, o escribir sobre voltajes y corrientes. ¡Haremos ambas cosas!

Espacio nulo de *A*: ¿Existe una combinación de columnas que proporcione $Ax = 0$? Normalmente, la respuesta se obtiene a partir de la eliminación, aunque aquí viene a primera vista. *La suma de las columnas es la columna cero.* El espacio nulo contiene a $x = (1, 1, 1, 1)$, ya que $Ax = 0$. La ecuación $Ax = b$ no tiene una solución única (en caso de tener alguna solución). Cualquier "vector constante" $x = (c, c, c, c)$ puede sumarse a cualquier solución particular de $Ax = b$. La solución completa tiene esta constante arbitraria c (como la $+C$ cuando se integra en cálculo).

Esto tiene un significado si se piensa que x_1, x_2, x_3, x_4 son ***los potenciales*** (voltajes) ***en los nodos***. Las cinco componentes de Ax proporcionan las ***diferencias*** de potencial a través de las cinco aristas. La diferencia a través de la arista 1 es $x_2 - x_1$, proveniente del ± 1 en el primer renglón.

La ecuación $Ax = b$ solicita: Dadas las diferencias $b_1, \ldots, b_5$, encontrar los potenciales verdaderos $x_1, \ldots, x_4$. ¡Pero esto es imposible de hacer! Todos los potenciales pueden aumentarse o disminuirse por la misma constante c, y las diferencias permanecen sin cambio, lo cual confirma que $x = (c, c, c, c)$ está en el espacio nulo de A. Estos son los únicos vectores en el espacio nulo, ya que $Ax = 0$ significa igual potencial a través de cada arista. El espacio nulo de esta matriz de incidencia es unidimensional. ***El rango es 4 − 1 = 3.***

Espacio columna: ¿Para cuáles diferencias $b_1, \ldots, b_5$ es posible resolver $Ax = b$? Para encontrar una prueba directa, considere la matriz. El renglón 1 más el renglón 3 es igual al renglón 2. En el miembro derecho se requiere $b_1 + b_3 = b_2$, o ninguna solución es posible. De manera semejante, el renglón 3 más el renglón 5 es igual al renglón 4. El miembro derecho debe satisfacer $b_3 + b_5 = b_4$, para que la eliminación llegue a $0 = 0$. Repitiendo, si b está en el espacio columna, entonces

$$b_1 - b_2 + b_3 = 0 \quad \text{y} \quad b_3 - b_4 + b_5 = 0. \tag{1}$$

Continuando la investigación, también se encuentra que los renglones 1 + 4 son iguales a los renglones 2 + 5. Pero esto no es nada nuevo: al restar las ecuaciones en (1) en realidad se obtiene $b_1 + b_4 = b_2 + b_5$. Hay *dos condiciones* sobre las cinco componentes, ya que la dimensión del espacio columna es $5 - 2$. Estas condiciones provienen de la eliminación, pero aquí tienen un significado sobre la gráfica.

Circuitos: La ley del voltaje de Kirchhoff establece que *la suma de las diferencias de potencial alrededor de un circuito debe ser cero*. Alrededor del circuito superior en la figura 2.6, las diferencias satisfacen $(x_2 - x_1) + (x_3 - x_2) = (x_3 - x_1)$. Estas diferencias son $b_1 + b_3 = b_2$. Para recorrer todo el circuito y regresar al mismo potencial, se requiere $b_3 + b_5 = b_4$.

2R La prueba para que b esté en el espacio columna es la ***ley del voltaje de Kirchhoff: La suma de las diferencias de potencial alrededor de un circuito debe ser cero.***

Espacio nulo izquierdo: Para resolver $A^{T}y = 0$, se encuentra su significado en la gráfica. El vector y tiene cinco componentes, una para cada arista. Estos números representan **corrientes** que circulan a lo largo de las cinco aristas. Debido a que A^{T} es de 4 por 5, las ecuaciones $A^{T}y = 0$ proporcionan cuatro condiciones sobre estas cinco corrientes. Se trata de condiciones de "conservación" en cada nodo: **En cada nodo, el flujo de entrada es igual al flujo de salida:**

$$A^{T}y = \mathbf{0} \qquad \begin{aligned} -y_1 - y_2 \qquad\quad &= 0 \quad \text{La corriente total hacia el nodo 1 es cero} \\ y_1 - y_3 - y_4 &= 0 \quad \text{hacia el nodo 2} \\ y_2 + y_3 - y_5 &= 0 \quad \text{hacia el nodo 3} \\ y_4 + y_5 &= 0 \quad \text{hacia el nodo 4} \end{aligned}$$

La belleza de la teoría de redes es que tanto A como A^{T} tienen papeles importantes.

Resolver $A^{T}y = 0$ significa encontrar un conjunto de corrientes que no se "amontone" en ningún nodo. El tráfico continúa circulando, y las soluciones más sencillas son **las corrientes alrededor de circuitos pequeños.** Nuestra gráfica tiene dos circuitos, y alrededor de cada circuito se envía 1 amp de corriente:

$$\textbf{Vectores de circuito} \quad y_1^{T} = \begin{bmatrix} 1 & -1 & 1 & 0 & 0 \end{bmatrix} \quad \text{y} \quad y_2^{T} = \begin{bmatrix} 0 & 0 & 1 & -1 & 1 \end{bmatrix}.$$

Cada circuito produce un vector y en el espacio nulo izquierdo. La componente $+1$ o -1 indica si la corriente va en el sentido de la flecha o en contra. Las combinaciones de y_1 y y_2 llenan el espacio nulo izquierdo, por lo que y_1 y y_2 son una base (la dimensión tendría que ser $m - r = 5 - 3 = 2$). De hecho, $y_1 - y_2 = (1, -1, 0, 1, -1)$ proporciona el gran circuito alrededor de la parte externa de la gráfica.

El espacio columna y el espacio nulo izquierdo están relacionados estrechamente. El espacio nulo izquierdo contiene a $y_1 = (1, -1, 1, 0, 0)$, y los vectores en el espacio columna satisfacen $b_1 - b_2 + b_3 = 0$. Así, $y^{T}b = 0$: ¡los vectores en el espacio columna y en el espacio nulo izquierdo son perpendiculares! Esto pronto se convertirá en la parte dos del "Teorema fundamental del álgebra lineal."

Espacio renglón: El espacio renglón de A contiene vectores en $\mathbf{R}^4$, pero no a todos los vectores. Su dimensión es el rango $r = 3$. Con la eliminación se encuentran tres renglones independientes, y también es posible ver la gráfica. Los tres primeros renglones son *dependientes* (renglón 1 + renglón 3 = renglón 2, y estas aristas forman un circuito). *Los renglones* 1, 2, 4 *son independientes porque las aristas* 1, 2, 4 *no contienen circuitos.*

Los renglones 1, 2, 4 son una base del espacio renglón. *En cada renglón, la suma de los elementos es cero.* Toda combinación (f_1, f_2, f_3, f_4) en el espacio renglón tiene la misma propiedad:

$$\boldsymbol{f} \textbf{ en el espacio renglón } f_1 + f_2 + f_3 + f_4 = 0 \quad \boldsymbol{x} \textbf{ en el espacio nulo } x = c(1, 1, 1, 1) \qquad (2)$$

Nuevamente, esto ilustra el teorema fundamental: el espacio renglón es perpendicular al espacio nulo. *Si f está en el espacio renglón y x está en el espacio nulo, entonces* $f^{T}x = 0$.

Para A^{T}, la ley básica de la teoría de redes es la ***ley de la corriente de Kirchhoff***. *El flujo total hacia cada nodo es cero.* Los números f_1, f_2, f_3, f_4 son fuentes de corriente hacia los nodos. La fuente f_1 debe equilibrar a $-y_1 - y_2$, que es el flujo que sale del nodo 1 (a lo largo de las aristas 1 y 2). Esta es la primera ecuación en $A^{T}y = f$. De manera semejante en los otros tres nodos, la conservación de la carga requiere *flujo de entrada = flujo de salida*. Lo hermoso es que A^{T} ***es exactamente la matriz derecha de la ley de la corriente***.

2S Las ecuaciones $A^{T}y = f$ en los nodos expresan la ***ley de la corriente de Kirchhoff***:

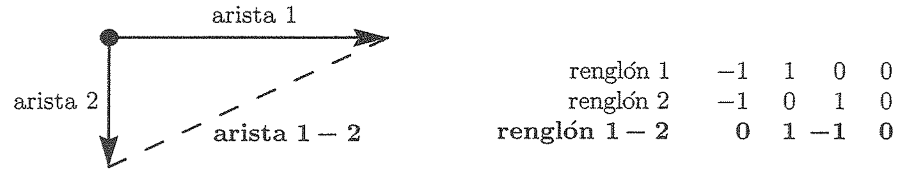
La corriente neta hacia cada nodo es cero. Flujo de entrada = flujo de salida.

Esta ley sólo puede cumplirse si la corriente total proveniente de fuera es $f_1 + f_2 + f_3 + f_4 = 0$. Con $f = 0$, la ley $A^{T}y = 0$ es satisfecha por ***una corriente que recorre el circuito.***

Árboles generadores y renglones independientes

Toda componente de y_1 y y_2 en el espacio nulo izquierdo es 1 o -1 o 0 (provenientes de los flujos en el circuito). ¡Lo mismo es cierto para $x = (1, 1, 1, 1)$ en el espacio nulo, y todos los elementos en $PA = LDU$! La cuestión clave es que todo paso de la eliminación tiene un significado para la gráfica.

El significado puede observarse en el primer paso de la eliminación para la matriz A: *reste el renglón* 1 *del renglón* 2. Esto sustituye a la arista 2 por una nueva arista "1 menos 2";



Ese paso de la eliminación destruye una arista y crea una nueva arista. Aquí la arista "1 − 2" es justo la arista 3 anterior en dirección opuesta. El siguiente paso de la eliminación produce ceros en el renglón 3 de la matriz. Esto demuestra que los renglones 1, 2, 3 son dependientes. *Los renglones son dependientes si las aristas correspondientes contienen un circuito.*

Al final de la eliminación se tiene un conjunto completo de r renglones independientes. **Estas r aristas constituyen un árbol: una gráfica sin circuitos.** Nuestra gráfica tiene $r = 3$, y las aristas 1, 2, 4 forman un árbol posible. El nombre completo es *árbol generador* porque el árbol "genera" todos los nodos de la gráfica. Un árbol generador tiene $n - 1$ aristas si la gráfica es conexa, y si se incluye una arista más se obtiene un circuito.

En el lenguaje del álgebra lineal, $n - 1$ es el rango de la matriz de incidencia A. La dimensión del espacio renglón es $n - 1$. El árbol generador que se obtiene de la eliminación constituye una base para el espacio renglón: cada arista del árbol corresponde a un renglón en la base.

El teorema fundamental del álgebra lineal relaciona las dimensiones de los subespacios:

Espacio nulo: dimensión 1, contiene a $x = (1, \ldots, 1)$.
Espacio columna: dimensión $r = n - 1$, cualesquiera $n - 1$ columnas son independientes.
Espacio renglón: dimensión $r = n - 1$, renglones independientes de cualquier árbol generador.
Espacio nulo izquierdo: dimensión $m - r = m - n + 1$, contiene ys de los circuitos.

Estos cuatro espacios proporcionan la ***fórmula de Euler***, que de alguna manera es el primer teorema en topología. Cuenta aristas de dimensión cero menos aristas unidimensionales más circuitos bidimensionales. Ahora cuenta con una demostración del álgebra lineal para cualquier gráfica conexa:

$$(\text{\# de nodos}) - (\text{\# de aristas}) + (\text{\# de circuitos}) = (n) - (m) + (m - n + 1) = \mathbf{1}. \qquad (3)$$

Para un simple circuito de 10 nodos y 10 aristas, el número de Euler es $10 - 10 + 1$. Si todos y cada uno de esos 10 nodos están unidos a un undécimo nodo en el centro, entonces $11 - 20 + 10$ sigue siendo 1.

Todo vector f en el espacio columna tiene $x^{\mathrm{T}}f = f_1 + \cdots + f_n = 0$; la suma de las corrientes provenientes del exterior es cero. Todo vector b en el espacio columna tiene $y^{\mathrm{T}}b = 0$; la suma de las diferencias de potencial es cero alrededor de todos los circuitos. En algún momento, x y y se vincularán mediante una tercera ley (***ley de Ohm para cada resistor***). Primero nos quedamos con la matriz A para presentar una aplicación que parece frívola pero no lo es.

Clasificación de los equipos de fútbol

Al final de esta temporada, los equipos de fútbol colegial son clasificados según varias encuestas. La clasificación es esencialmente un promedio de opiniones, y algunas veces se vuelve vaga después de las primeras doce universidades. Se desea clasificar a todos los equipos de acuerdo con una base más matemática.

El primer paso es reconocer la gráfica. Si el equipo j jugó contra el equipo k, entonces entre ambos hay una arista. Los *equipos* son los *nodos*, y los *juegos* son las *aristas*. Hay unos cientos de nodos y algunos miles de aristas, a las que se asignará una dirección mediante una flecha que va del equipo visitante al equipo local. En la figura 2.7 se muestra parte de la Liga Ivy, así como algunos equipos serios, y también una universidad que no es famosa por sus logros en fútbol colegial. Afortunadamente para esa universidad (en donde estoy escribiendo estas líneas), la gráfica no es conexa. En términos matemáticos, no es posible demostrar que el MIT no es el número 1 (a menos que ocurra que juegue contra alguien).

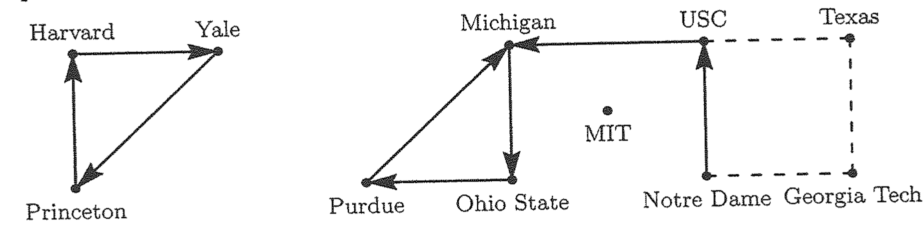


Figura 2.7 Parte de la gráfica de fútbol americano colegial.

Si el fútbol fuese perfectamente consistente, a cada equipo podría asignarse un "potencial" x_j. Así, si un equipo visitante v juega con el equipo local h, entonces ganaría el equipo que tuviera mayor potencial. En el caso ideal, la diferencia b en el resultado sería exactamente igual a la diferencia $x_h - x_v$ en sus potenciales. ¡Ni siquiera tendrían que jugar! Habría acuerdo por completo en que el mejor equipo es aquel con mayor potencial.

Este método presenta dos dificultades (por lo menos). Se está intentando encontrar un número x para cada equipo, y se desea $x_h - x_v = b_i$ para cada juego. Esto significa unos cuantos miles de ecuaciones y sólo unos cuantos cientos de incógnitas. Las ecuaciones $x_h - x_v = b_i$ van a un sistema lineal $Ax = b$, donde A es una ***matriz de incidencia***. Todo juego tiene un renglón, con $+1$ en la columna h, y -1 en la columna v, con la finalidad de indicar qué equipos jugaron ese partido.

Primera dificultad: Si b no está en el espacio columna no hay solución. Los resultados deben ajustarse perfectamente o no es posible encontrar potenciales exactos. Segunda dificultad: Si A tiene vectores diferentes de cero en su espacio nulo, los potenciales x no están bien determinados. En el primer caso x no existe; en el segundo, x no es único. Quizá estén presentes ambas dificultades.

El espacio nulo siempre contiene el vector de 1s, ya que A sólo ve las *diferencias* $x_h - x_v$. Para determinar los potenciales, arbitrariamente podría asignarse potencial cero a Harvard. (Estoy hablando matemáticamente, no asumiéndolo como un hecho). Pero si la gráfica no es conexa, entonces toda pieza por separado de la gráfica contribuye a un vector en el espacio nulo. Incluso está el vector con $x_{\mathrm{MIT}} = 1$ y todos los demás $x_j = 0$. Es necesario unir no sólo a Harvard sino a un equipo en cada pieza. (No hay nada de injusto en asignar potencial cero; si todos los demás potenciales están abajo de cero, entonces el equipo unido se clasifica primero.) La dimensión del espacio nulo es *el número de piezas* de la gráfica: y no habrá forma de clasificar una pieza contra otra, ya que éstas no juegan partidos.

Parece que el espacio columna es más difícil de describir. ¿Cuáles resultados se ajustan perfectamente bien a un conjunto de potenciales? Ciertamente $Ax = b$ es irresoluble si Harvard derrota a Yale, Yale derrota a Princeton y Princeton derrota a Harvard. Más que lo anterior, la suma de las diferencias en los resultados de ese circuito de juegos ***debe ser cero***.

Ley de Kirchhoff para diferencias en los resultados $b_{\mathrm{HY}} + b_{\mathrm{YP}} + b_{\mathrm{PH}} = 0$.

Ésta también es una ley del álgebra lineal. $Ax = b$ puede resolverse cuando b satisface las mismas dependencias lineales que los renglones de A. Así, la eliminación lleva a $0 = 0$.

En realidad, casi con toda certeza b no está en el espacio columna. Los resultados de fútbol no son tan consistentes. Para obtener una clasificación es posible utilizar ***mínimos cuadrados***: hacer a Ax lo más próximo posible de b. Eso se verá en el capítulo 3, y sólo se menciona un ajuste. El ganador obtiene un bono de 50 e incluso 100 puntos por arriba de la diferencia en resultados. En caso contrario, ganar por 1 está demasiado próximo a perder por 1. Esto hace que las clasificaciones calculadas se aproximen bastante a las encuestas, y el doctor Leake (Notre Dame) proporcionó un análisis completo en *Management Science in Sports* (1976).

Después de escribir esta subsección, encontré lo siguiente en el *New York Times*:

> En sus clasificaciones finales en 1985, la computadora ubicó a Miami (10-2) en séptimo lugar, arriba de Tennessee (9-1-2). Pocos días después de su publicación, al departamento de deportes del *Times* empezaron a llegar paquetes con naranjas y cartas de enojo, enviados por los fanáticos descontentos de Tennessee. La irritación surgió del hecho de que Tennessee apabulló a Miami 35-7 en el tazón del azúcar. Las encuestas finales de AP y UPI clasificaron en cuarto lugar a Tennessee, con Miami bastante más abajo.
>
> Ayer en la mañana llegaron nueve envases de naranjas al muelle de carga. Fueron enviadas al hospital Bellevue con una advertencia de que la calidad y contenido de las naranjas era incierto.

Tanto, para esta aplicación del álgebra lineal.

Redes y matemáticas discretas aplicadas

Una gráfica se vuelve una ***red*** cuando a las aristas se asignan números $c_1, \ldots, c_m$. El número c_i puede ser la *longitud* de la arista i, o su *capacidad*, o su *rigidez* (si contiene a un resorte), o su *conductancia* (si contiene un resistor). Estos números van en una matriz diagonal C que es de m por m. C refleja "propiedades materiales", en contraste con la matriz de incidencia A, que proporciona información de las conexiones.

Nuestra descripción será en términos de electricidad. Sobre la arista i, la conductancia es c_i y la resistencia es $1/c_i$. La ley de Ohm establece que la corriente y_i que pasa

por el resistor es proporcional a la caída de tensión e_i:

Ley de Ohm $y_i = c_i e_i$ (corriente) = (conductancia)(caída de tensión).

Lo anterior también se escribe $E = IR$, caída de tensión igual a la corriente multiplicada por la resistencia. Como una ecuación vectorial sobre todas las aristas a la vez, la ***ley de Ohm es*** $y = Ce$.

Para completar el marco de referencia se requieren la ley del voltaje y la ley de la corriente de Kirchhoff:

LVK: La suma de las caídas de tensión alrededor de cada circuito es cero.
LCK: La suma de las corrientes y_i (y f_i) hacia cada nodo es cero.

La ley del voltaje permite asignar potenciales $x_1, \ldots, x_n$ a los nodos. Luego, las diferencias alrededor de un circuito proporcionan una sumatoria como $(x_2 - x_1) + (x_3 - x_2) + (x_1 - x_3) = 0$, en la que todo se cancela. La ley de la corriente pide sumar las corrientes hacia cada nodo por la multiplicación de $A^{\mathrm{T}}y$. Si no hay fuentes de corriente externas, *la ley de la corriente de Kirchhoff es* $A^{\mathrm{T}}y = 0$.

La otra ecuación es la ley de Ohm, pero es necesario encontrar la caída de tensión e a través del resistor. Con la multiplicación Ax se obtuvo la diferencia de potencial entre los nodos. Al invertir los signos, $-Ax$ proporciona *la caída* en potencial. Parte de esta caída puede deberse a una ***batería*** en la arista de intensidad b_i. El resto de la caída es $e = b - Ax$ a través del resistor:

$$\textbf{Ley de Ohm} \qquad y = C(b - Ax) \ \text{ o bien } \ C^{-1}y + Ax = b. \tag{4}$$

Las ***leyes fundamentales de equilibrio*** combinan las leyes de Ohm y Kirchhoff en un problema central de las matemáticas aplicadas. Estas ecuaciones aparecen en todas partes:

$$\textbf{Ecuaciones de equilibrio} \qquad \begin{aligned} C^{-1}y + Ax &= b \\ A^{\mathrm{T}}y &= f. \end{aligned} \tag{5}$$

Este es un sistema lineal simétrico, del cual ha desaparecido e. Las incógnitas son las corrientes y y los potenciales x. Usted ve la matriz simétrica por bloques:

$$\textbf{Forma de bloques} \qquad \begin{bmatrix} C^{-1} & A \\ A^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = \begin{bmatrix} b \\ f \end{bmatrix}. \tag{6}$$

Para eliminación por bloques el pivote es C^{-1}, el multiplicador es $A^{\mathrm{T}}C$, y la sustracción manda a A^{T} abajo del pivote. El resultado es

$$\begin{bmatrix} C^{-1} & A \\ 0 & -A^{\mathrm{T}}CA \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = \begin{bmatrix} b \\ f - A^{\mathrm{T}}Cb \end{bmatrix}$$

La ecuación sólo para x está en el renglón inferior, con la matriz simétrica $A^{\mathrm{T}}CA$:

$$\textbf{Ecuación fundamental} \qquad A^{\mathrm{T}}CAx = A^{\mathrm{T}}Cb - f. \tag{7}$$

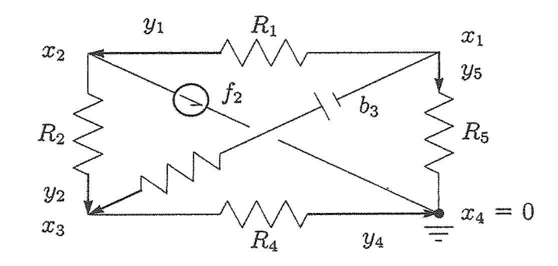
Luego, la sustitución hacia atrás en la primera ecuación produce y. Nada misterioso: se sustituye $y = C(b - Ax)$ en $A^{\mathrm{T}}y = f$ para obtener (7).

Observación importante Un potencial debe fijarse de antemano: $x_n = 0$. El n-ésimo nodo está ***conectado a tierra***, y la n-ésima columna de la matriz de incidencia original se ha

eliminado. La matriz resultante es lo que ahora se entiende por A, sus $n - 1$ columnas son independientes. La matriz cuadrada $A^{\mathrm{T}}CA$, que es la clave para resolver la ecuación (7) para x, es una matriz invertible de orden $n - 1$:

$$\underset{n-1 \text{ por } m}{\begin{bmatrix} & A^{\mathrm{T}} & \end{bmatrix}} \underset{m \text{ por } m}{\begin{bmatrix} \\ C \\ \\ \end{bmatrix}} \underset{m \text{ por } n-1}{\begin{bmatrix} \\ A \\ \\ \end{bmatrix}} = \underset{n-1 \text{ por } n-1}{\begin{bmatrix} A^{\mathrm{T}}CA \end{bmatrix}}$$

Ejemplo 1 Suponga que una batería b_3 y una fuente de corriente f_2 (y cinco resistores) conectan cuatro nodos. El nodo 4 está conectado a tierra y el potencial $x_4 = 0$ es fijo.



La primera cuestión es la ley de la corriente $A^{\mathrm{T}}y = f$ en los nodos 1, 2, 3:

$$\begin{aligned} -y_1 - y_3 - y_5 &= 0 \\ y_1 - y_2 &= f_2 \\ y_2 + y_3 - y_4 &= 0 \end{aligned} \quad \text{tiene} \quad A^{\mathrm{T}} = \begin{bmatrix} -1 & 0 & -1 & 0 & -1 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 & 0 \end{bmatrix}.$$

Para el nodo 4 no se ha escrito ninguna ecuación, donde la ley de la corriente es $y_4 + y_5 + f_2 = 0$. Esto se concluye al haber sumado las otras tres ecuaciones.

La otra ecuación es $C^{-1}y + Ax = b$. Los potenciales x están conectados a las corrientes y por la ley de Ohm. La matriz diagonal C contiene las cinco conductancias $c_i = 1/R_i$. El miembro derecho explica la batería de intensidad b_3 en la arista 3. La forma de bloque tiene a $C^{-1}y + Ax = b$ arriba de $A^{\mathrm{T}}y = f$:

$$\begin{bmatrix} C^{-1} & A \\ A^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = \begin{bmatrix} R_1 & & & & & -1 & 1 & 0 \\ & R_2 & & & & 0 & -1 & 1 \\ & & R_3 & & & -1 & 0 & 1 \\ & & & R_4 & & 0 & 0 & -1 \\ & & & & R_5 & -1 & 0 & 0 \\ -1 & 0 & -1 & 0 & -1 & & & \\ 1 & -1 & 0 & 0 & 0 & & & \\ 0 & 1 & 1 & -1 & 0 & & & \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ b_3 \\ 0 \\ 0 \\ 0 \\ f_2 \\ 0 \end{bmatrix}$$

El sistema es de 8 por 8, con cinco corrientes y cuatro potenciales. La eliminación de las ys reduce al sistema de 3 por 3 $A^{\mathrm{T}}CAx = A^{\mathrm{T}}Cb - f$. La matriz $A^{\mathrm{T}}CA$ contiene los recíprocos $c_i = 1/R_i$ (porque en la eliminación se dividen los pivotes). También se muestran el

cuarto renglón y la cuarta columna, provenientes del nodo conectado a tierra, fuera de la matriz de 3 por 3:

$$A^{\mathrm{T}}CA = \left[\begin{array}{ccc|c} c_1+c_3+c_5 & -c_1 & -c_3 & -c_5 \\ -c_1 & c_1+c_2 & -c_2 & 0 \\ -c_3 & -c_2 & c_2+c_3+c_4 & -c_4 \\ \hline -c_5 & 0 & -c_4 & c_4+c_5 \end{array}\right] \begin{array}{l} \text{(nodo 1)} \\ \text{(nodo 2)} \\ \text{(nodo 3)} \\ \text{(nodo 4)} \end{array}$$

El primer elemento es $1 + 1 + 1$, o $c_1 + c_3 + c_5$ cuando se incluye C, ya que las aristas 1, 3, 5 tocan el nodo 1. El siguiente elemento diagonal es $1 + 1$ o $c_1 + c_2$, provenientes de las aristas que tocan al nodo 2. Fuera de la diagonal, las cs aparecen con signo menos. *Las aristas hacia el nodo* 4 conectado a tierra pertenecen al *cuarto* renglón y a la cuarta columna, que se borran una vez que la columna 4 se elimina de A (haciendo invertible a $A^{\mathrm{T}-}CA$). La matriz de 4 por 4 debe cumplir que la suma de todos los renglones y todas las columnas es cero, y que (1, 1, 1, 1) debe estar en el espacio nulo.

Observe que $A^{\mathrm{T}}CA$ es simétrica. Tiene pivotes positivos y proviene del **marco de referencia básico de las matemáticas aplicadas** que se ilustra en la figura 2.8.

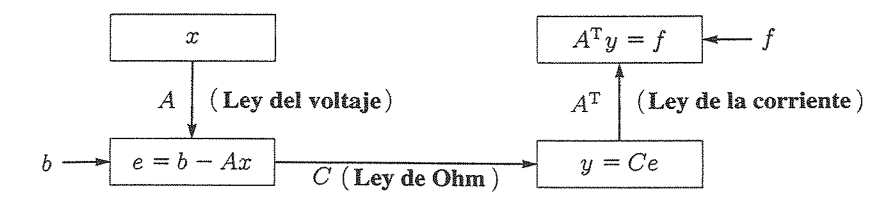


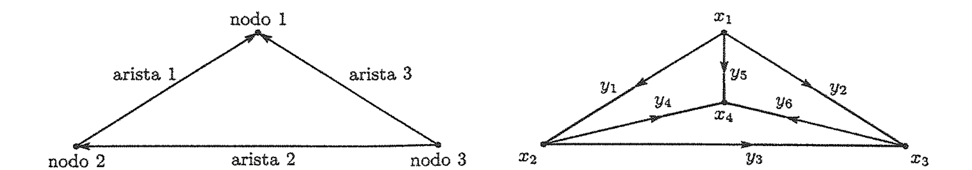
Figura 2.8 Marco de referencia para equilibrio: fuentes b y f, tres pasos para $A^{\mathrm{T}}CA$.

En mecánica, x y y se vuelven desplazamientos y esfuerzos. En fluidos, las incógnitas son la presión y el caudal de flujo. En estadística, e es el error y x es el mejor ajuste por mínimos cuadrados a los datos. Estas ecuaciones matriciales y las ecuaciones diferenciales correspondientes se encuentran en nuestro libro de texto *Introduction to Applied Mathematics*, y en el nuevo *Applied Mathematics and Scientific Computing*. (Consulte la página `www.wellesleycambridge.com`.)

Este capítulo termina en un punto culminante: el ***planteamiento*** de un problema fundamental en matemáticas aplicadas. A menudo para esto se requiere más habilidad que para la ***solución*** del problema. En el capítulo 1 se resolvieron ecuaciones lineales como primer paso del álgebra lineal. Para plantear las ecuaciones se requiere el conocimiento más profundo del capítulo 2. La contribución de las matemáticas, y de la gente, no es la computación sino la inteligencia.

Conjunto de problemas 2.5

1. Para la gráfica triangular de 3 nodos de la siguiente figura, escriba la matriz de incidencia A de 3 por 3. Encuentre una solución de $Ax = 0$, y describa todos los otros vectores en el espacio nulo de A. Encuentre una solución de $A^{\mathrm{T}}y = 0$, y describa todos los otros vectores en el espacio nulo izquierdo de A.
2. Para la misma matriz de 3 por 3, demuestre directamente a partir de las columnas que todo vector b en el espacio columna satisface $b_1 + b_2 - b_3 = 0$. Deduzca el mismo



hecho a partir de los tres renglones: las ecuaciones en el sistema $Ax = b$. ¿Qué significa esto sobre las diferencias de potencial alrededor de un circuito?

3. Demuestre directamente a partir de los renglones que todo vector f en el espacio renglón satisface $f_1 + f_2 + f_3 = 0$. Deduzca el mismo hecho a partir de las tres ecuaciones $A^{\mathrm{T}}y = f$. ¿Qué significa esto cuando las fs son corrientes hacia los nodos?
4. Calcule la matriz $A^{\mathrm{T}}A$ de 3 por 3, y demuestre que es simétrica aunque singular: ¿qué vectores están en su espacio nulo? Al eliminar la última columna de A (y también el último renglón de A^{T}) se queda con la matriz de 2 por 2 en la esquina superior izquierda; demuestre que es *no* singular.
5. Escriba la matriz diagonal C con elementos c_1, c_2, c_3 en medio, y calcule $A^{\mathrm{T}}CA$. Demuestre nuevamente que la matriz de 2 por 2 en la esquina superior izquierda es invertible.
6. Escriba la matriz de incidencia A de 6 por 4 para la segunda gráfica en la figura. El vector (1, 1, 1, 1) está en el espacio nulo de A, pero ahora ahí hay $m - n + 1 = 3$ vectores independientes que satisfacen $A^{\mathrm{T}}y = 0$. Encuentre tres vectores y, y *únalos con los circuitos en la gráfica*.
7. Si esa segunda gráfica representa seis juegos entre cuatro equipos, y las diferencias en los resultados son $b_1, \ldots, b_6$, ¿cuándo es posible asignar potenciales $x_1, \ldots, x_4$ de modo que la diferencia de potenciales coincida con las bs? Usted está encontrando (a partir de las leyes de Kirchhoff o por eliminación) las condiciones que hacen resoluble a $Ax = b$.
8. Escriba las dimensiones de los cuatro subespacios fundamentales para esta matriz de incidencia de 6 por 4, y una base para cada subespacio.
9. Calcule $A^{\mathrm{T}}A$ y $A^{\mathrm{T}}CA$, donde la matriz diagonal C de 6 por 6 tiene los elementos $c_1, \ldots, c_6$. ¿Cómo puede afirmar a partir de la gráfica dónde aparecen las cs en la diagonal principal de $A^{\mathrm{T}}CA$?
10. Trace una gráfica con aristas numeradas y dirigidas (y con nodos numerados) cuya matriz de incidencia es

$$A = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

Esta gráfica, ¿es un árbol? (los renglones de A, ¿son independientes?) Demuestre que al eliminar la última arista se obtiene un árbol generador. Así, los renglones restantes son una base de ____.

11. Sin la última columna de la A precedente, y con los números 1, 2, 2, 1 en la diagonal de C, escriba el sistema de 7 por 7

$$\begin{aligned} C^{-1}y + Ax &= 0 \\ A^{\mathrm{T}}y \qquad &= f. \end{aligned}$$

Al eliminar y_1, y_2, y_3, y_4 se queda con tres ecuaciones $A^{\mathrm{T}}CAx = -f$ para x_1, x_2, x_3. Resuelva las ecuaciones cuando $f = (1, 1, 6)$. Con estas corrientes entrando en los nodos 1, 2, 3 de la red, ¿cuáles son los potenciales en los nodos y las corrientes en las aristas?

12. Si A es una matriz de incidencia de 12 por 7 de una gráfica conexa, ¿cuál es su rango? ¿Cuántas variables libres hay en la solución de $Ax = b$? ¿Cuántas variables libres hay en la solución de $A^{\mathrm{T}}y = f$? ¿Cuántas aristas deben eliminarse para obtener un árbol generador?

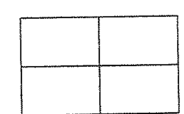
13. En la gráfica anterior con 4 nodos y 6 aristas, encuentre todos los 16 árboles generadores.

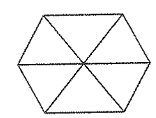
14. Si MIT derrota a Harvard por 35 a 0, Yale empata con Harvard, y Princeton derrota a Yale por 7 a 6, ¿qué diferencias en los resultados en los otros tres juegos (H-P, MIT-P, MIT-Y) permiten diferencias de potencial que coincidan con las diferencias en los resultados? Si las diferencias en los resultados se conocen en un árbol generador, entonces se conocen para todos los juegos.

15. En nuestro método de clasificación de fútbol, ¿es necesario considerar la fuerza de la oposición, o ya está considerada?

16. Si entre cada par de nodos hay una arista (una gráfica completa), ¿cuántas aristas hay? La gráfica tiene n nodos, y no se permiten aristas de un nodo hacia sí mismo.

17. Para las dos gráficas que se muestran a continuación, compruebe la *fórmula de Euler*: **(# de nodos)** − **(# de aristas)** + **(# de circuitos)** = 1.





18. Multiplique matrices para encontrar $A^{\mathrm{T}}A$, y conjeture cómo provienen sus elementos de la gráfica:
   a) La diagonal de $A^{\mathrm{T}}A$ indica cuántos (cuántas) ____ hay en cada nodo.
   b) Los elementos −1 o 0 fuera de la diagonal indican cuáles pares de nodos son ____.

19. ¿Por qué el espacio nulo de $A^{\mathrm{T}}A$ contiene a (1, 1, 1, 1)? ¿Cuál es su rango?

20. ¿Por qué una gráfica con $n = 6$ nodos tiene $m = 15$ aristas? Un árbol generador que une a todos los seis nodos tiene ____ aristas. ¡Hay $n^{n-2} = 6^4$ árboles generadores!

21. La *matriz de adyacencia* de una gráfica tiene $M_{ij} = 1$ si los nodos i y j están unidos por una arista (en caso contrario $M_{ij} = 0$). Para la gráfica del problema 6 con 6 nodos y 4 aristas, escriba M y también M^2. ¿Por qué $(M^2)_{ij}$ cuenta el número de *rutas de* 2 *pasos* del nodo i al nodo j?

2.6 TRANSFORMACIONES LINEALES

Se sabe cómo una matriz mueve los subespacios alrededor cuando se multiplica por A. El espacio nulo se va al vector cero. Todos los vectores van hacia el espacio columna, ya que Ax siempre es una combinación de las columnas. Pronto verá algo hermoso: que A lleva su espacio renglón en su espacio columna, y que sobre estos espacios de dimensión r es 100% invertible. Esta es la verdadera acción de A. Parcialmente está escondida por los espacios nulos y los espacios nulos izquierdos, que están a ángulos rectos y llevan su propio camino (hacia cero).

Lo que importa ahora es lo que ocurre *dentro* del espacio, lo cual significa dentro del espacio n-dimensional, si A es de n por n. Esto requiere un análisis más detallado.

Suponga que x es un vector n-dimensional. Cuando A se multiplica por x, ***transforma*** ese vector en uno nuevo Ax. Esto ocurre en todo punto x del espacio n-dimensional $\mathbf{R}^n$. La matriz A transforma todo el espacio, o lo "mapea en sí mismo". En la figura 2.9 se ilustran cuatro transformaciones que provienen de estas matrices:

$$A = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix}$$

1. Un múltiplo de la matriz identidad, $A = cI$, ***alarga*** cada vector por el mismo factor c. Todo el espacio se dilata o contrae (o de alguna forma pasa por el origen y sale por el lado opuesto, cuando c es negativo).

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

2. Una matriz ***rotación*** hace girar a todo el espacio alrededor del origen. Este ejemplo gira 90° a todos los vectores, transformando cada punto (x, y) en $(-y, x)$.

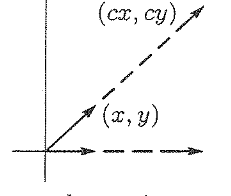
$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

3. Una matriz ***reflexión*** transforma a todos los vectores en su imagen especular. En este ejemplo el espejo es la recta $x = y$ a 45°, y un punto como (2, 2) permanece sin cambio. Un punto como (2, −2) se invierte en (−2, 2). Sobre una combinación como $v = (2, 2) + (2, -2) = (4, 0)$, la matriz preserva una parte e invierte la otra. El resultado es $Av = (2, -2) + (-2, 2) = (0, 4)$.

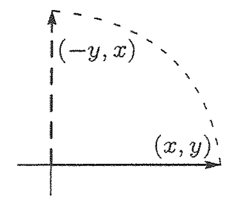
   ¡Esa matriz reflexión es también una matriz permutación! Algebraicamente es tan sencillo, mandar (x, y) a (y, x), que la representación geométrica se ocultó.

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

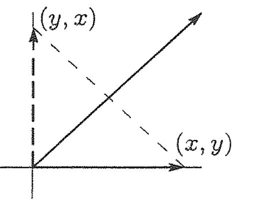
4. Una matriz ***proyección*** transforma todo el espacio en un subespacio de dimensión inferior (no invertible). El ejemplo transforma cada vector (x, y) en el plano hasta el punto más próximo $(x, 0)$ sobre el eje horizontal. Este eje es el espacio columna de A. El eje y que proyecta a (0, 0) es el espacio nulo.



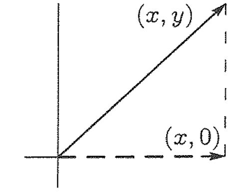
alargamiento



rotación de 90°



reflexión (espejo a 45°)



proyección sobre el eje

Figura 2.9 Transformaciones del plano realizadas por cuatro matrices.

Estos ejemplos pudieron presentarse en tres dimensiones. Hay matrices para alargar la Tierra, hacerla girar o reflejarla a través del plano del ecuador (el Polo Norte se transforma en el Polo Sur). Hay una matriz que proyecta todo sobre ese plano (ambos polos al centro). También es importante reconocer que las matrices no pueden hacer todo, y que algunas transformaciones $T(x)$ *no son posibles* con Ax:

i) Es imposible mover el origen, ya que $A0 = 0$ para toda matriz.
ii) Si el vector x se va en x', entonces $2x$ debe irse en $2x'$. En general, cx debe irse en cx', ya que $A(cx) = c(Ax)$.
iii) Si los vectores x y y se van en x' y y', entonces su suma $x + y$ se debe ir en $x' + y'$, ya que $A(x + y) = Ax + Ay$.

La multiplicación de matrices impone estas reglas sobre la transformación. La segunda regla contiene a la primera (tome $c = 0$ para obtener $A0 = 0$). La regla iii) se vio en acción cuando (4, 0) se reflejó a través de la recta a 45°. Se separó en (2, 2) + (2, −2) y ambas partes fueron reflejadas por separado. Lo mismo puede hacerse para las proyecciones: separar, proyectar por separado, y sumar las proyecciones. Estas reglas son válidas para ***cualquier transformación que provenga de una matriz.***

Esta importancia les ha ganado una denominación: las transformaciones que cumplen las reglas i) a iii) se denominan ***transformaciones lineales***. Las reglas pueden combinarse en un requerimiento:

2T Para todos los números c y d y todos los vectores x y y, la multiplicación de matrices satisface la regla de linealidad:

$$A(cx + dy) = c(Ax) + d(Ay). \tag{1}$$

Toda transformación $T(x)$ que cumple este requerimiento es una ***transformación lineal***.

Cualquier matriz lleva de inmediato a una transformación lineal. La pregunta más interesante es en la dirección opuesta: *¿Toda transformación lineal lleva a una matriz*? El objetivo de esta sección es encontrar la respuesta: *sí*. Este es el fundamento de un método del álgebra lineal —empezar con la propiedad 1) y desarrollar sus consecuencias—; esto es mucho más abstracto que el método más importante de este libro. Aquí se eligió empezar directamente con matrices, y a continuación se verá cómo representan transformaciones lineales.

Una transformación no necesita ir de $\mathbf{R}^n$ al mismo espacio $\mathbf{R}^n$. Se permite absolutamente transformar vectores en $\mathbf{R}^n$ en vectores en otro espacio distinto $\mathbf{R}^m$. ¡Esto es exactamente lo que hace una matriz de m por n! El vector original x tiene n componentes, y el vector transformado Ax tiene m componentes. La regla de linealidad la cumplen igualmente las matrices rectangulares, de modo que éstas también producen transformaciones lineales.

A estas alturas del libro, no hay razón para detenerse. Las operaciones sobre la condición de linealidad 1) son la suma y la multiplicación por un escalar, aunque x y y no necesitan ser vectores columna en $\mathbf{R}^n$. Éstos no son los únicos espacios. Por definición, *cualquier espacio vectorial permite las combinaciones $cx + dy$*, los "vectores" son x y y, aunque en realidad pueden ser polinomios, matrices o funciones $x(t)$ y $y(t)$. En la medida en que la transformación cumpla la ecuación 1), es lineal.

Como ejemplos se toman los espacios $\mathbf{P}_n$, donde los vectores son polinomios $p(t)$ de grado n. Se denotan como $p = a_0 + a_1t + \cdots + a_nt^n$ y la dimensión del espacio vectorial es $n + 1$ (debido a que con el término constante, hay $n + 1$ coeficientes).

Ejemplo 1 La operación de ***diferenciación***, $A = d/dt$, es lineal:

$$Ap(t) = \frac{d}{dt}(a_0 + a_1t + \cdots + a_nt^n) = a_1 + \cdots + na_nt^{n-1}. \tag{2}$$

El espacio nulo de esta A es el espacio unidimensional de constantes: $da_0/dt = 0$. El espacio columna es el espacio n-dimensional $\mathbf{P}_{n-1}$; el miembro derecho de la ecuación (2) siempre está en ese espacio. La suma de la nulidad (= 1) y el rango (= n) es la dimensión del espacio original $\mathbf{P}_n$.

Ejemplo 2 La ***integración*** desde 0 hasta t también es lineal (*lleva* $\mathbf{P}_n$ *a* $\mathbf{P}_{n+1}$):

$$Ap(t) = \int_0^t (a_0 + \cdots + a_nt^n)\,dt = a_0t + \cdots + \frac{a_n}{n+1}t^{n+1}. \tag{3}$$

Esta vez no hay espacio nulo (¡excepto por el vector cero, como siempre!) pero la integración no produce todos los polinomios en $\mathbf{P}_{n+1}$. El miembro derecho de la ecuación (3) no tiene término constante. Probablemente los polinomios constantes sean el espacio nulo izquierdo.

Ejemplo 3 La ***multiplicación*** por un polinomio fijo como $2 + 3t$ es lineal:

$$Ap(t) = (2 + 3t)(a_0 + \cdots + a_nt^n) = 2a_0 + \cdots + 3a_nt^{n+1}.$$

De nuevo, esto transforma $\mathbf{P}_n$ en $\mathbf{P}_{n+1}$, sin espacio nulo excepto $p = 0$.

En estos ejemplos (y en casi todos los ejemplos), no es difícil comprobar la linealidad. Incluso, difícilmente parece interesante hacerlo. Si hay linealidad, prácticamente es imposible ignorarla. A pesar de ello, se trata de la propiedad más importante que puede tener una transformación.* Por supuesto, la mayor parte de las transformaciones no son lineales: por ejemplo, para elevar al cuadrado al polinomio ($Ap = p^2$), o sumar 1 ($Ap = p + 1$), o preservar los coeficientes positivos ($A(t - t^2) = t$). Son las transformaciones lineales, y *sólo éstas*, lo que lleva a las matrices.

Transformaciones representadas por matrices

La linealidad tiene una consecuencia crucial: ***Si se conoce Ax para cada vector en una base, entonces se conoce Ax para cada vector en todo el espacio.*** Suponga que esta base consta de los n vectores $x_1, \ldots, x_n$. Cualquier otro vector x es una combinación de estos vectores particulares (generan el espacio). Así, la linealidad determina Ax:

$$\textbf{Linealidad si } x = c_1x_1 + \cdots + c_nx_n \text{ entonces } Ax = c_1(Ax_1) + \cdots + c_n(Ax_n). \tag{4}$$

Una vez que la transformación $T(x) = Ax$ ha decidido qué hacer con los vectores de la base, ya no tiene libertad disponible. El resto es determinado por la linealidad. El requerimiento 1) para dos vectores x y y lleva a la condición 4) para n vectores $x_1, \ldots, x_n$. La transformación tiene manos libres con los vectores en la base (son independientes). Una vez que éstos se establecen, también se establece la transformación de cada vector.

*En orden de importancia, quizá en segundo lugar esté la invertibilidad.

Ejemplo 4 ¿Qué transformación lineal lleva x_1 y x_2 a Ax_1 y Ax_2?

$$x_1 = \begin{bmatrix}1\\0\end{bmatrix} \text{ se va en } Ax_1 = \begin{bmatrix}2\\3\\4\end{bmatrix}; \qquad x_2 = \begin{bmatrix}0\\1\end{bmatrix} \text{ se va en } Ax_2 = \begin{bmatrix}4\\6\\8\end{bmatrix}.$$

Debe ser la multiplicación $T(x) = Ax$ por la matriz

$$A = \begin{bmatrix}2 & 4\\3 & 6\\4 & 8\end{bmatrix}.$$

Empezando con una base distinta (1, 1) y (2, −1), esta misma A también es la única transformación lineal con

$$A\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}6\\9\\12\end{bmatrix} \quad \text{y} \quad A\begin{bmatrix}2\\-1\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$

Luego se encuentran matrices que representan diferenciación e integración. ***Primero debe decidirse sobre una base***. Para los polinomios de grado 3 hay una elección natural para los cuatro vectores de la base:

$$\textbf{Base de } \mathbf{P}_3 \qquad p_1 = 1, \quad p_2 = t, \quad p_3 = t^2, \quad p_4 = t^3.$$

Esta base no es única (nunca lo es), pero se requiere de una elección, y ésta es la más conveniente. Las derivadas de estos cuatro vectores de la base son 0, 1, $2t$, $3t^2$:

$$\textbf{Acción de } d/dt \quad Ap_1 = 0, \quad Ap_2 = p_1, \quad Ap_3 = 2p_2, \quad Ap_4 = 3p_3. \tag{5}$$

"d/dt" actúa de inmediato exactamente como una matriz, pero ¿cuál matriz? Suponga que se está trabajando en el espacio tetradimensional de costumbre con la base de costumbre, los vectores de coordenadas $p_1 = (1, 0, 0, 0)$, $p_2 = (0, 1, 0, 0)$, $p_3 = (0, 0, 1, 0)$, $p_4 = (0, 0, 0, 1)$. La matriz es decidida por la ecuación (5):

$$\textbf{Matriz diferenciación} \qquad A_{\text{dif}} = \begin{bmatrix}0 & 1 & 0 & 0\\0 & 0 & 2 & 0\\0 & 0 & 0 & 3\\0 & 0 & 0 & 0\end{bmatrix}.$$

Ap_1 es su primera columna, que es cero. Ap_2 es la segunda columna, que es p_1. Ap_3 es $2p_2$, y Ap_4 es $3p_3$. El espacio nulo contiene a p_1 (la derivada de una constante es cero). El espacio columna contiene a p_1, p_2, p_3 (la derivada de una cúbica es una cuadrática). La derivada de una combinación como $p = 2 + t - t^2 - t^3$ es decidida por la linealidad, y no hay nada nuevo en eso: es la forma en que todos diferenciamos. Sería absurdo memorizar la derivada de cada polinomio.

La matriz puede diferenciar este $p(t)$, ¡porque las matrices incluyen linealidad!

$$\frac{dp}{dt} = Ap \rightarrow \begin{bmatrix}0 & 1 & 0 & 0\\0 & 0 & 2 & 0\\0 & 0 & 0 & 3\\0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}2\\1\\-1\\-1\end{bmatrix} = \begin{bmatrix}1\\-2\\-3\\0\end{bmatrix} \rightarrow 1 - 2t - 3t^2.$$

En breve, *la matriz contiene toda la información esencial*. Si se conoce la base, y se conoce la matriz, entonces se conoce la transformación de cada vector.

La codificación de la información es sencilla. Para transformar un espacio en sí mismo, basta una base. Una transformación de un espacio en otro requiere una base para cada espacio.

2U Suponga que los vectores $x_1, \ldots, x_n$ son una base del espacio **V**, y que los vectores $y_1, \ldots, y_m$ son una base de **W**. Cada transformación lineal T de **V** a **W** es representada por una matriz A. La j-ésima columna se encuentra aplicando T al j-ésimo vector x_j de la base, y escribiendo $T(x_j)$ como una combinación de las ys:

$$\textbf{Columna } j \textbf{ de } A \qquad T(x_j) = Ax_j = a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m. \tag{6}$$

Para la matriz diferenciación, la columna 1 provino del primer vector de la base, $p_1 = 1$. Su derivada es cero, de modo que la columna 1 es cero. La última columna provino de $(d/dt)t^3 = 3t^2$. Debido a que $3t^2 = 0p_1 + 0p_2 + 3p_3 + 0p_4$, la última columna contenía a 0, 0, 3, 0. La regla (6) construye la matriz, columna por columna.

Para la integración se hace lo mismo. Ahí se va de cúbicas a cuárticas, transformando $\mathbf{V} = \mathbf{P}_3$ en $\mathbf{W} = \mathbf{P}_4$, por lo que se requiere una base de **W**. La elección natural es $y_1 = 1$, $y_2 = t$, $y_3 = t^2$, $y_4 = t^3$, $y_5 = t^4$, generando los polinomios de grado 4. La matriz A es de m por n, o de 5 por 4. Proviene de la aplicación de la integración a cada vector de la base de **V**:

$$\int_0^t 1\,dt = t \quad \text{o bien,} \quad Ax_1 = y_2, \quad \ldots, \quad \int_0^t t^3\,dt = \frac{1}{4}t^4 \quad \text{o bien,} \quad Ax_4 = \frac{1}{4}y_5.$$

$$\textbf{Matriz integración} \qquad A_{\text{int}} = \begin{bmatrix}0 & 0 & 0 & 0\\1 & 0 & 0 & 0\\0 & \frac{1}{2} & 0 & 0\\0 & 0 & \frac{1}{3} & 0\\0 & 0 & 0 & \frac{1}{4}\end{bmatrix}.$$

La diferenciación y la integración son *operaciones inversas*. O por lo menos la integración *seguida* de la diferenciación regresa a la función original. Para hacer que esto ocurra para las matrices, se requiere la matriz diferenciación de cuárticas a cúbicas, que es de 4 por 5:

$$A_{\text{dif}} = \begin{bmatrix}0 & 1 & 0 & 0 & 0\\0 & 0 & 2 & 0 & 0\\0 & 0 & 0 & 3 & 0\\0 & 0 & 0 & 0 & 4\end{bmatrix} \quad \text{y} \quad A_{\text{dif}}A_{\text{int}} = \begin{bmatrix}1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1\end{bmatrix}.$$

La diferenciación es una ***inversa izquierda*** de la integración. ¡Las matrices rectangulares no pueden tener inversas por ambos lados! En el orden opuesto, $A_{\text{int}}A_{\text{dif}} = I$ no puede ser cierto. El producto de 5 por 5 tiene ceros en su columna 1. La derivada de una constante es cero. En las otras columnas $A_{\text{int}}A_{\text{dif}}$ es la identidad, y la integral de la derivada de t^n es t^n.

Rotaciones Q, proyecciones P y reflexiones H

Esta sección empieza con rotaciones de 90°, proyecciones sobre el eje x, y reflexiones a través de la recta a 45°. Sus matrices son especialmente sencillas:

$$Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad H = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

(rotación) (proyección) (reflexión)

Las transformaciones lineales subyacentes del plano x-y también son sencillas. Sin embargo, las rotaciones a través de otros ángulos, las proyecciones sobre otras rectas, y las reflexiones en otros espejos son casi tan fáciles de visualizar. Siguen siendo transformaciones lineales, suponiendo que el origen esté fijo: $A0 = 0$. *Deben* estar representadas por matrices. Usando la base natural $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ y $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$, se desea descubrir estas matrices.

1. **Rotación** En la figura 2.10 se muestra la rotación de un ángulo θ. También muestra el efecto sobre los dos vectores de la base. La primera va a (cos θ, sen θ), cuya longitud sigue siendo 1; está en la "recta θ". El segundo vector base (0, 1) rota en (−sen θ, cos θ). Por la regla (6), estos números van a las columnas de la matriz (se usa c para cos θ y s para sen θ). Esta familia de rotaciones Q_θ constituye una oportunidad perfecta para probar la correspondencia entre las transformaciones y las matrices:

*¿La **inversa** de Q_θ es igual a $Q_{-\theta}$* (rotación hacia atrás por un ángulo θ)? *Sí.*

$$Q_\theta Q_{-\theta} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} c & s \\ -s & c \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

*¿El **cuadrado** de Q_θ es igual a $Q_{2\theta}$* (rotación por un ángulo doble)? *Sí.*

$$Q_\theta^2 = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} = \begin{bmatrix} c^2 - s^2 & -2cs \\ 2cs & c^2 - s^2 \end{bmatrix} = \begin{bmatrix} \cos 2\theta & -\operatorname{sen} 2\theta \\ \operatorname{sen} 2\theta & \cos 2\theta \end{bmatrix}.$$

*¿El **producto** de Q_θ y Q_φ es igual a $Q_{\theta+\varphi}$* (rotación por un ángulo θ y luego φ)? *Sí.*

$$Q_\theta Q_\varphi = \begin{bmatrix} \cos\theta\cos\varphi - \operatorname{sen}\theta\operatorname{sen}\varphi & \cdots \\ \operatorname{sen}\theta\cos\varphi + \cos\theta\operatorname{sen}\varphi & \cdots \end{bmatrix} = \begin{bmatrix} \cos(\theta+\varphi) & \cdots \\ \operatorname{sen}(\theta+\varphi) & \cdots \end{bmatrix}.$$

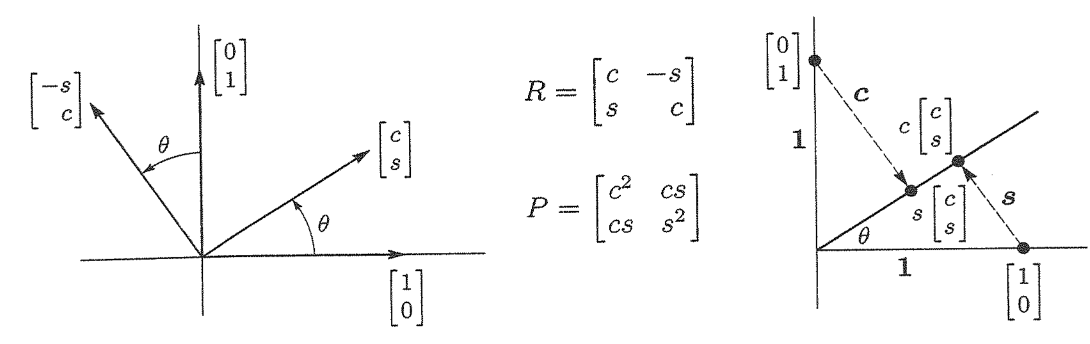


Figura 2.10 Rotación en un ángulo θ (izquierda). Proyección sobre la recta θ (derecha).

El último caso contiene a los dos primeros. La inversa aparece cuando φ es $-\theta$, y el cuadrado aparece cuando φ es $+\theta$. Estas tres ecuaciones fueron decididas por identidades trigonométricas (y constituyen una nueva manera de recordar estas identidades). No es accidental que la respuesta a las tres preguntas sea sí. *La multiplicación de matrices se define exactamente de modo que **el producto de las matrices corresponda al producto de las transformaciones.***

2V Suponga que A y B son transformaciones lineales de $\mathbf{V}$ a $\mathbf{W}$ y de $\mathbf{U}$ a $\mathbf{V}$. Su producto AB empieza con un vector u en $\mathbf{U}$, va a Bu en $\mathbf{V}$, y termina con ABu en $\mathbf{W}$. Esta "composición" AB de nuevo es una transformación lineal (de $\mathbf{U}$ a $\mathbf{W}$). Su matriz es el producto de las matrices individuales que representan A y B.

Para $A_{\text{dif}}A_{\text{int}}$, la transformación compuesta era la identidad (y $A_{\text{int}}A_{\text{dif}}$ aniquilaron a todas las constantes). Para rotaciones, el orden de la multiplicación no importa. Así, $\mathbf{U} = \mathbf{V} = \mathbf{W}$ es el plano x-y, y $Q_\theta Q_\varphi$ es lo mismo que $Q_\varphi Q_\theta$. Para una rotación y una reflexión, el orden sí es importante.

Nota técnica: Para construir las matrices, se requieren bases de $\mathbf{V}$ y $\mathbf{W}$, y luego para $\mathbf{U}$ y $\mathbf{V}$. Al mantener la misma base para $\mathbf{V}$, el producto de matrices va correctamente de la base en $\mathbf{U}$ a la base en $\mathbf{W}$. Si la transformación A se distingue de su matriz (sea ésta $[A]$), entonces la regla del producto 2V se vuelve extremadamente concisa: $[AB] = [A][B]$. La regla para multiplicar matrices, en el capítulo 1 estaba totalmente determinada por este requerimiento: debe corresponder al producto de transformaciones lineales.

2. **Proyección** En la figura 2.10 también se muestra la proyección de (1, 0) sobre la recta θ. La longitud de la proyección es $c = \cos\theta$. Observe que *el punto* de proyección no es (c, s), como yo consideraba erróneamente; la longitud de ese vector es 1 (es la rotación), de modo que es necesario multiplicar por c. De manera semejante, la longitud de la proyección de (0, 1) es s, y cae en $s(c, s) = (cs, s^2)$. Así se obtiene la segunda columna de la matriz proyección P:

$$\textbf{Proyección sobre la recta } \boldsymbol{\theta} \qquad P = \begin{bmatrix} c^2 & cs \\ cs & s^2 \end{bmatrix}.$$

Esta matriz no tiene inversa, porque la transformación no tiene inversa. Los puntos sobre la recta perpendicular son proyectados sobre el origen; esa recta es el espacio nulo de P. ¡Los puntos sobre la recta θ son proyectados sobre sí mismos! Proyectar dos veces es lo mismo que proyectar una vez, y $P^2 = P$:

$$P^2 = \begin{bmatrix} c^2 & cs \\ cs & s^2 \end{bmatrix}^2 = \begin{bmatrix} c^2(c^2+s^2) & cs(c^2+s^2) \\ cs(c^2+s^2) & s^2(c^2+s^2) \end{bmatrix} = P.$$

Por supuesto, $c^2 + s^2 = \cos^2\theta + \operatorname{sen}^2\theta = 1$. ***Una matriz proyección es igual a su propio cuadrado.***

3. **Reflexión** En la figura 2.11 se muestra la reflexión de (1, 0) en la recta θ. La longitud de la reflexión es igual a la longitud del original, como era después de la rotación; sin embargo, en este caso la recta θ permanece donde está. La recta perpendicular invierte la dirección; todos los puntos pasan directamente a través del espejo. La linealidad decide el resto.

$$\textbf{Matriz reflexión} \qquad H = \begin{bmatrix} 2c^2 - 1 & 2cs \\ 2cs & 2s^2 - 1 \end{bmatrix}.$$

$$2c\begin{bmatrix}c\\s\end{bmatrix}-\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}2c^2-1\\2cs\end{bmatrix}$$

$$H = 2P - I = \begin{bmatrix}2c^2-1 & 2cs\\ 2cs & 2s^2-1\end{bmatrix}$$

imagen + original = 2 × proyección

$$Hx + x = 2Px$$

$\begin{bmatrix}0\\1\end{bmatrix}$ $\begin{bmatrix}1\\0\end{bmatrix}$ θ

$$2s\begin{bmatrix}c\\s\end{bmatrix}-\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}2cs\\2s^2-1\end{bmatrix}$$

Figura 2.11 Reflexión a través de la recta θ: geometría y matriz.

La matriz H posee la extraordinaria propiedad $H^2 = I$. ***Dos reflexiones devuelven el original.*** Una reflexión es su propia inversa, $H = H^{-1}$, lo cual resulta evidente a partir de la geometría pero es menos evidente a partir de la matriz. Un método es a través de la relación de las reflexiones con las proyecciones: $H = 2P - I$. Esto significa que $Hx + x = 2Px$: la imagen más el original es igual a dos veces la proyección. También confirma que $H^2 = I$:

$$H^2 = (2P - I)^2 = 4P^2 - 4P + I = I, \quad \text{ya que } P^2 = P.$$

Otras transformaciones Ax pueden incrementar la longitud de x; el alargamiento y el esfuerzo cortante se encuentran en los ejercicios. Cada ejemplo tiene una matriz que lo representa, lo cual constituye la cuestión más importante de esta sección. Sin embargo, también está la cuestión de elegir una base, y se recalca que *la matriz depende de la elección de la base*. Suponga que el primer vector en la base está ***sobre la recta*** θ y que el segundo vector en la base es ***perpendicular***:

i) La matriz proyección es regresada por $P = \begin{bmatrix}1&0\\0&0\end{bmatrix}$. Esta matriz se construye como siempre: su primera columna proviene del primer vector en la base (proyectado en sí mismo). La segunda columna proviene del vector en la base que es proyectado en cero.

ii) Para reflexiones, esa misma base proporciona $H = \begin{bmatrix}1&0\\0&-1\end{bmatrix}$. El segundo vector en la base es reflejado sobre su negativo para producir esta segunda columna. La matriz H sigue siendo $2P - I$ cuando la misma base se usa para H y P.

iii) Para rotaciones, la matriz no cambia. Estas rectas siguen rotando un ángulo θ, y $Q = \begin{bmatrix}c&-s\\s&c\end{bmatrix}$ como antes.

Toda la cuestión de elegir la mejor base es absolutamente fundamental, por lo que en el capítulo 5 se abordará nuevamente. El objetivo es hacer diagonal a la matriz, como se logró para P y H. Para hacer diagonal a Q se requieren vectores complejos, ya que todos los vectores reales están rotados.

Aquí se menciona el efecto que tiene sobre la matriz un *cambio de base*, mientras las transformaciones lineales permanecen igual. ***La matriz A** (o Q, o P, o H) **se modifica a*** $S^{-1}AS$. Así, una simple transformación es representada por varias matrices (mediante bases diferentes, tomadas en cuenta por S). La teoría de los vectores característicos conduce a la fórmula $S^{-1}AS$, así como a la mejor base.

Conjunto de problemas 2.6

1. ¿Qué matriz tiene el efecto de rotar cada vector un ángulo de 90° y luego proyectar el resultado sobre el eje x? ¿Qué matriz representa la proyección sobre el eje x seguida de la proyección sobre el eje y?

2. El producto de 5 reflexiones y 8 rotaciones del plano x-y, ¿produce una rotación o una reflexión?

3. La matriz $A = \begin{bmatrix}2&0\\0&1\end{bmatrix}$ produce un ***alargamiento*** en la dirección x. Trace el círculo $x^2 + y^2 = 1$ y a su alrededor trace los puntos $(2x, y)$ que resultan de la multiplicación por A. ¿Qué forma tiene esa curva?

4. Toda recta sigue siendo recta después de una transformación lineal. Si z está a la mitad entre x y y, demuestre que Az está a la mitad entre Ax y Ay.

5. La matriz $A = \begin{bmatrix}1&0\\3&1\end{bmatrix}$ produce una transformación por ***esfuerzo cortante***, que deja el eje y sin cambio. Bosqueje este efecto en el eje x, indicando lo que ocurre a $(1, 0)$ y $(2, 0)$ y $(-1, 0)$, y cómo se transforma todo el eje.

6. ¿Cuáles son las matrices de 3 por 3 que representan las transformaciones que

a) proyectan cada vector sobre el plano x-y?

b) reflejan cada vector a través del plano x-y?

c) rotan el plano x-y un ángulo de 90°, dejando sólo al eje z?

d) rotan un ángulo de 90° al plano x-y, luego al plano x-z, y luego al plano y-z?

e) realizan las tres rotaciones, pero cada una de un ángulo de 180°?

7. En el espacio $\mathbf{P}_3$ de polinomios cúbicos, ¿qué matriz representa d^2/dt^2? Construya la matriz de 4 por 4 a partir de la base estándar $1, t, t^2, t^3$. Encuentre su espacio nulo y su espacio columna. ¿Qué significan éstos en términos de polinomios?

8. De los cúbicos $\mathbf{P}_3$ hasta los polinomios de cuarto grado $\mathbf{P}_4$, ¿qué matriz representa la multiplicación por $2 + 3t$? Las columnas de la matriz A de 5 por 4 provienen de la aplicación de la transformación a $1, t, t^2, t^3$.

9. Las soluciones de la ecuación diferencial $d^2u/dt^2 = u$ forman un espacio vectorial (ya que las combinaciones de soluciones siguen siendo soluciones). Encuentre dos soluciones independientes, con la finalidad de obtener una base para ese espacio solución.

10. Con valores iniciales $u = x$ y $du/dt = y$ en $t = 0$, ¿qué combinación de los vectores en la base del problema 9 resuelve $u'' = u$? Esta transformación de valores iniciales a solución es lineal. ¿Cuál es su matriz de 2 por 2 (usando $x = 1, y = 0$ y $x = 0, y = 1$ como base de $\mathbf{V}$, y su base de $\mathbf{W}$)?

11. Compruebe directamente a partir de $c^2 + s^2 = 1$ que las matrices reflexión satisfacen $H^2 = I$.

12. Suponga que A es una transformación lineal del plano x-y en sí mismo. ¿Qué hace $A^{-1}(x + y) = A^{-1}x + A^{-1}y$? Si A está representada por la matriz M, explique por qué A^{-1} está representada por M^{-1}.

13. El producto $(AB)C$ de transformaciones lineales empieza con un vector x y produce $u = Cx$. Luego, la regla 2V aplica AB a u y llega a $(AB)Cx$.

a) ¿Este resultado es el mismo si se aplican por separado C, luego B y por último A?
b) ¿Este resultado es el mismo si se aplica BC seguido de A? Los paréntesis son innecesarios y la ley asociativa $(AB)C = A(BC)$ se cumple para transformaciones lineales. Esta es la mejor demostración de la misma ley para matrices.

14. Demuestre que T^2 es una transformación lineal si T es lineal (de $\mathbf{R}^3$ a $\mathbf{R}^3$).

15. El espacio de todas las matrices de 2 por 2 tiene los cuatro "vectores" en la base

$$\begin{bmatrix}1&0\\0&0\end{bmatrix},\quad \begin{bmatrix}0&1\\0&0\end{bmatrix},\quad \begin{bmatrix}0&0\\1&0\end{bmatrix},\quad \begin{bmatrix}0&0\\0&1\end{bmatrix}.$$

Para la transformación lineal de *trasposición*, encuentre su matriz A respecto a esta base. ¿Por qué $A^2 = I$?

16. Encuentre la matriz permutación cíclica de 4 por 4: (x_1, x_2, x_3, x_4) se transforma en $Ax = (x_2, x_3, x_4, x_1)$. ¿Cuál es el efecto de A^2? Demuestre que $A^3 = A^{-1}$.

17. Encuentre la matriz A de 4 por 3 que representa un *desplazamiento derecho*: (x_1, x_2, x_3) se transforma en $(0, x_1, x_2, x_3)$. También encuentre la matriz *desplazamiento izquierdo* B de $\mathbf{R}^4$ de regreso a $\mathbf{R}^3$, transformando (x_1, x_2, x_3, x_4) en (x_2, x_3, x_4). ¿Cuáles son los productos AB y BA?

18. En el espacio vectorial P_3 de todos los $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3$, sea $\mathbf{S}$ el subconjunto de los polinomios con $\int_0^1 p(x)\,dx = 0$. Compruebe que $\mathbf{S}$ es un subespacio y encuentre una base.

19. Una transformación *no lineal* es invertible si $T(x) = b$ tiene exactamente una solución para toda b. El ejemplo si $T(x) = x^2$ no es invertible porque $x^2 = b$ tiene dos soluciones para b positiva y ninguna solución para b negativa. ¿Cuáles de las siguientes transformaciones (de los números reales $\mathbf{R}^1$ a los números reales $\mathbf{R}^1$) son invertibles? Ninguna es lineal, ni siquiera *c*).

a) $T(x) = x^3$. *b)* $T(x) = e^x$.
c) $T(x) = x + 11$. *d)* $T(x) = \cos x$.

20. ¿Cuál es el eje y el ángulo de rotación para la transformación que lleva (x_1, x_2, x_3) a (x_2, x_3, x_1)?

21. Una transformación lineal debe dejar fijo al vector cero: $T(0) = 0$. Demuestre esto a partir de $T(v + w) = T(v) + T(w)$ escogiendo $w =$ ____. También demuéstrelo a partir del requerimiento $T(cv) = cT(v)$ escogiendo $c =$ ____.

22. ¿Cuáles de las siguientes transformaciones no son lineales? La entrada es $v = (v_1, v_2)$.

a) $T(v) = (v_2, v_1)$. *b)* $T(v) = (v_1, v_1)$.
c) $T(v) = (0, v_1)$. *d)* $T(v) = (0, 1)$.

23. Si S y T son lineales con $S(v) = T(v) = v$, entonces ¿$S(T(v)) = v$ o v^2?

24. Suponga que $T(v) = v$, excepto que $T(0, v_2) = (0, 0)$. Demuestre que esta transformación satisface $T(cv) = cT(v)$ pero no $T(v + w) = T(v) + T(w)$.

25. ¿Cuáles de las siguientes transformaciones satisfacen $T(v + w) = T(v) + T(w)$, y cuáles satisfacen $T(cv) = cT(v)$?

a) $T(v) = v/\|v\|$. *b)* $T(v) = v_1 + v_2 + v_3$.
c) $T(v) = (v_1, 2v_2, 3v_3)$. *d)* $T(v) =$ la mayor componente de v.

26. Para las siguientes transformaciones de $\mathbf{V} = \mathbf{R}^2$ en $\mathbf{W} = \mathbf{R}^2$, encuentre $T(T(v))$.

a) $T(v) = -v$.
b) $T(v) = v + (1, 1)$.
c) $T(v) = 90°$ rotación $= (-v_2, v_1)$.
d) $T(v) =$ proyección $= \left(\frac{v_1 + v_2}{2}, \frac{v_1 + v_2}{2}\right)$.

27. La transformación "cíclica" T se define por $T(v_1, v_2, v_3) = (v_2, v_3, v_1)$. ¿Cuál es $T(T(T(v)))$? ¿Cuál es $T^{100}(v)$?

28. Encuentre el *contradominio* y el *kernel* (estos términos son nuevos para designar el espacio columna y el espacio nulo) de T.

a) $T(v_1, v_2) = (v_2, v_1)$. *b)* $T(v_1, v_2, v_3) = (v_1, v_2)$.
c) $T(v_1, v_2) = (0, 0)$. *d)* $T(v_1, v_2) = (v_1, v_1)$.

29. Una transformación lineal de $\mathbf{V}$ a $\mathbf{W}$ tiene una *inversa* de $\mathbf{W}$ a $\mathbf{V}$ cuando el contradominio es todo $\mathbf{W}$ y el kernel sólo contiene a $v = 0$. ¿Por qué estas transformaciones no son invertibles?

a) $T(v_1, v_2) = (v_2, v_2)$ $\mathbf{W} = \mathbf{R}^2$.
b) $T(v_1, v_2) = (v_1, v_2, v_1 + v_2)$ $\mathbf{W} = \mathbf{R}^3$.
c) $T(v_1, v_2) = v_1$ $\mathbf{W} = \mathbf{R}^1$.

30. Suponga que una T lineal transforma $(1, 1)$ en $(2, 2)$ y $(2, 0)$ en $(0, 0)$. Encuentre $T(v)$ cuando

a) $v = (2, 2)$. *b)* $v = (3, 1)$. *c)* $v = (-1, 1)$. *d)* $v = (a, b)$.

Los problemas 31 a 35 pueden ser más difíciles. El espacio de entrada V contiene a todas las matrices *M* de 2 por 2.

31. M es cualquier matriz de 2 por 2 y $A = \begin{bmatrix}1&2\\3&4\end{bmatrix}$. La transformación lineal T se define como $T(M) = AM$. ¿Cuáles reglas de la multiplicación de matrices muestran que T es lineal?

32. Suponga que $A = \begin{bmatrix}1&2\\3&6\end{bmatrix}$. Demuestre que la matriz identidad I no está en el contradominio de T. Encuentre una matriz M diferente de cero tal que $T(M) = AM$ es cero.

33. Suponga que T transpone a toda matriz M. Intente encontrar una matriz A que haga $AM = M^{\mathrm{T}}$ para toda M. Demuestre que ninguna matriz A puede hacer lo anterior. *Para los profesores*: ¿Esta es una transformación lineal que no proviene de una matriz?

34. La transformación T que transpone a toda matriz es definitivamente lineal. ¿Cuáles de las siguientes propiedades adicionales son ciertas?

a) $T^2 =$ transformación identidad.
b) El kernel de T es la matriz cero.
c) Todas las matrices están en el contradominio de T.
d) $T(M) = -M$ es imposible.

35. Suponga que $T(M) = \begin{bmatrix}1&0\\0&0\end{bmatrix}\begin{bmatrix}M\end{bmatrix}\begin{bmatrix}0&0\\0&1\end{bmatrix}$. Encuentre una matriz con $T(M) \neq 0$. Describa todas las matrices con $T(M) = 0$ (el kernel de T) y todas las matrices de salida $T(M)$ (el contradominio de T).

Los problemas 36 a 40 son sobre cambio de base.

36. *a)* ¿Qué matriz transforma (1, 0) en (2, 5) y (0, 1) en (1, 3)?
b) ¿Qué matriz transforma (2, 5) en (1, 0) y (1, 3) en (0, 1)?
c) ¿Por qué ninguna matriz transforma (2, 6) en (1, 0) y (1, 3) en (0, 1)?

37. *a)* ¿Qué matriz M transforma (1, 0) y (0, 1) en (r, t) y (s, u)?
b) ¿Qué matriz N transforma (a, c) y (b, d) en (1, 0) y (0, 1)?
c) ¿Qué condición sobre a, b, c, d hace imposible el inciso *b)*?

38. *a)* ¿Cómo M y N del problema 37 producen una matriz que transforma (a, c) en (r, t) y (b, d) en (s, u)?
b) ¿Qué matriz transforma (2, 5) en (1, 1) y (1, 3) en (0, 2)?

39. Si se conservan los mismos vectores en la base pero se escriben en otro orden, la matriz de cambio de base M es una matriz ____. Si los vectores en la base se preservan en orden pero se cambian sus longitudes, entonces M es una matriz ____.

40. La matriz que transforma (1, 0) y (0, 1) en (1, 4) y (1, 5) es $M =$ ____. La combinación $a(1, 4) + b(1, 5)$ igual a (1, 0) tiene $(a, b) = (\ ,\ )$. ¿Cómo se relacionan estas coordenadas de (1, 0) con M o M^{-1}?

41. ¿Cuáles son las tres ecuaciones para A, B, C si la parábola $Y = A + Bx + Cx^2$ es igual a 4 en $x = a$, 5 en $x = b$, y 6 en $x = c$? Encuentre el determinante de la matriz de 3 por 3. ¿Para qué números a, b, c es imposible encontrar esta parábola Y?

42. Suponga que v_1, v_2, v_3 son vectores característicos para T. Esto significa que $T(v_i) = \lambda_i v_i$ para $i = 1, 2, 3$. ¿Cuál es la matriz para T cuando las bases de entrada y de salida son las vs?

43. Toda transformación lineal invertible puede tener a I como su matriz. Para la base de salida se elige $w_i = T(v_i)$. ¿Por qué T debe ser invertible?

44. Suponga que T es una reflexión a través del eje x y que S es una reflexión a través del eje y. El dominio V es el plano x-y. Si $v = (x, y)$, ¿cuál es $S(T(v))$? Encuentre una descripción más sencilla del producto ST.

45. Suponga que T es una reflexión a través de la recta a 45°, y que S es una reflexión a través del eje y. Si $v = (2, 1)$, entonces $T(v) = (1, 2)$. Encuentre $S(T(v))$ y $T(S(v))$. Con esto se demuestra que en general $ST \neq TS$.

46. Demuestre que el producto ST de dos reflexiones es una rotación. Multiplique estas matrices reflexión para encontrar el ángulo de rotación:

$$\begin{bmatrix} \cos 2\theta & \operatorname{sen} 2\theta \\ \operatorname{sen} 2\theta & -\cos 2\theta \end{bmatrix} \quad \begin{bmatrix} \cos 2\alpha & \operatorname{sen} 2\alpha \\ \operatorname{sen} 2\alpha & -\cos 2\alpha \end{bmatrix}.$$

47. La *matriz de Hadamard* de 4 por 4 está integrada completamente por +1 y −1:

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

Encuentre H^{-1} y escriba $v = (7, 5, 3, 1)$ como una combinación de las columnas de H.

48. Suponga que se tienen dos bases $v_1, \ldots, v_n$ y $w_1, \ldots, w_n$ de $\mathbf{R}^n$. Si un vector tiene coeficientes b_i en una base y c_i en la otra base, ¿cuál es la matriz de cambio de base en $b = Mc$? Empiece con

$$b_1v_1 + \cdots + b_nv_n = Vb = c_1w_1 + \cdots + c_nw_n = Wc.$$

Su respuesta representa $T(v) = v$ con base de entrada de vs y base de salida de ws. Debido a que las bases son distintas, la matriz no es I.

49. ¿Falso o verdadero? Si se conoce $T(v)$ para n vectores diferentes de cero en R^n, entonces se conoce $T(v)$ para cada vector en R^n.

50. (Recomendado) Suponga que todos los vectores x en el cuadrado unitario $0 \leq x_1 \leq 1$, $0 \leq x_2 \leq 1$ se transforman en Ax (A es de 2 por 2).
a) ¿Cuál es la forma de la región transformada (toda Ax)?
b) ¿Para qué matrices A esta región es cuadrada?
c) ¿Para qué matrices A es una recta?
d) ¿Para qué matrices A la nueva área sigue siendo 1?

Capítulo 2 Ejercicios de repaso

2.1 Encuentre una base para los siguientes subespacios de $\mathbf{R}^4$:
a) Los vectores para los cuales $x_1 = 2x_4$.
b) Los vectores para los cuales $x_1 + x_2 + x_3 = 0$ y $x_3 + x_4 = 0$.
c) El subespacio generado por (1, 1, 1, 1), (1, 2, 3, 4), y (2, 3, 4, 5).

2.2 Proporcione una base para describir un subespacio bidimensional de $\mathbf{R}^3$ que no contenga ninguno de los vectores de coordenadas (1, 0, 0), (0, 1, 0), (0, 0, 1).

2.3 ¿Falso o verdadero? Proporcione un contraejemplo si es falso:
a) Si los vectores $x_1, \ldots, x_m$ generan un subespacio S, entonces $\dim S = m$.
b) La intersección de dos subespacios de un espacio vectorial no puede ser vacía.
c) Si $Ax = Ay$, entonces $x = y$.
d) El espacio renglón de A tiene una base única que puede calcularse reduciendo A a forma escalonada.
e) Si una matriz cuadrada A tiene columnas independientes, también A^2 tiene columnas independientes.

2.4 ¿Cuál es la forma escalonada U de A?

$$A = \begin{bmatrix} 1 & 2 & 0 & 2 & 1 \\ -1 & -2 & 1 & 1 & 0 \\ 1 & 2 & -3 & -7 & -2 \end{bmatrix}.$$

¿Cuáles son las dimensiones de sus cuatro subespacios fundamentales?

2.5 Encuentre el rango y el espacio nulo de

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 2 \\ 1 & 1 & 1 & 0 \end{bmatrix}.$$

2.6 Encuentre bases para los cuatro subespacios fundamentales asociados con

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

2.7 ¿Cuál es la solución más general de $u + v + w = 1$, $u - w = 2$?

2.8 *a*) Construya una matriz cuyo espacio nulo contenga al vector $x = (1, 1, 2)$.
b) Construya una matriz cuyo espacio nulo izquierdo contenga a $y = (1, 5)$.
c) Construya una matriz cuyo espacio columna sea generado por $(1, 1, 2)$ y cuyo espacio renglón sea generado por $(1, 5)$.
d) Si se tienen tres vectores cualesquiera en $\mathbf{R}^6$ y tres vectores cualesquiera en $\mathbf{R}^5$, ¿existe alguna matriz de 6 por 5 cuyo espacio columna es generado por los primeros tres y cuyo espacio renglón es generado por los segundos tres?

2.9 En el espacio vectorial de matrices de 2 por 2,
a) ¿el conjunto de matrices con rango 1 es un subespacio?
b) ¿qué subespacio es generado por las matrices permutación?
c) ¿qué subespacio es generado por las matrices positivas (todos los $a_{ij} > 0$)?
d) ¿qué subespacio es generado por las matrices invertibles?

2.10 Invente un espacio vectorial que contenga a todas las transformaciones lineales de $\mathbf{R}^n$ a $\mathbf{R}^n$. Defina una regla para la suma. ¿Cuál es la dimensión del espacio vectorial?

2.11 *a*) Encuentre el rango de A, y proporcione una base de su espacio nulo.

$$A = LU = \begin{bmatrix} 1 & & & \\ 2 & 1 & & \\ 2 & 1 & 1 & \\ 3 & 2 & 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

b) Los 3 primeros renglones de U son una base del espacio renglón de A: ¿falso o verdadero?
Las columnas 1, 3, 6 de U son una base del espacio columna de A: ¿falso o verdadero?
Los cuatro renglones de A son una base del espacio renglón de A: ¿falso o verdadero?
c) Encuentre tantos vectores b linealmente independientes como sea posible para los cuales $Ax = b$ tenga una solución.
d) En la eliminación sobre A, ¿qué múltiplo del tercer renglón se restó para eliminar el cuarto renglón?

2.12 Si A es una matriz de n por $n - 1$, y su rango es $n - 2$, ¿cuál es la dimensión de su espacio nulo?

2.13 Use eliminación para encontrar los factores triangulares en $A = LU$, si

$$A = \begin{bmatrix} a & a & a & a \\ a & b & b & b \\ a & b & c & c \\ a & b & c & d \end{bmatrix}.$$

¿Qué condiciones deben imponerse a los números a, b, c, d para que las columnas sean linealmente independientes?

2.14 ¿Los vectores $(1, 1, 3)$, $(2, 3, 6)$, y $(1, 4, 3)$ son una base de $\mathbf{R}^3$?

2.15 ¿Qué se sabe de $C(A)$ cuando el número de soluciones de $Ax = b$ es
a) 0 o 1, dependiendo de b.
b) ∞, independientemente de b.
c) 0 o ∞, dependiendo de b.
d) 1, sin importar b.

2.16 En el ejercicio anterior, ¿cómo está relacionado r con m y n en cada ejemplo?

2.17 Si x es un vector en $\mathbf{R}^n$, y $x^{\mathrm{T}}y = 0$ para toda y, demuestre que $x = 0$.

2.18 Si A es una matriz de n por n tal que $A^2 = A$ y rango $A = n$, demuestre que $A = I$.

2.19 ¿Qué subespacio de matrices de 3 por 3 es generado por las matrices elementales E_{ij}, con 1s en la diagonal y cuando mucho un elemento diferente de cero debajo de la diagonal?

2.20 ¿Cuántas matrices permutación de 5 por 5 hay? ¿Son linealmente independientes? ¿Generan el espacio de todas las matrices de 5 por 5? No es necesario que las escriba todas.

2.21 ¿Cuál es el rango de la matriz de n por n con todos los elementos igual a 1? ¿Qué puede decir sobre la "matriz de tablero de ajedrez", con $a_{ij} = 0$ cuando $i + j$ es par, $a_{ij} = 1$ cuando $i + j$ es impar?

2.22 *a*) ¿Qué condiciones deben imponerse a b para que $Ax = b$ tenga una solución, para las siguientes A y b?

$$A = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 2 & 4 & 0 & 1 \end{bmatrix} \quad \text{y} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

b) Encuentre una base para el espacio nulo de A.
c) Encuentre la solución general de $Ax = b$, cuando exista una solución.
d) Encuentre una base para el espacio columna de A.
e) ¿Cuál es el rango de A^{T}?

2.23 ¿Cómo puede construirse una matriz que transforme los vectores de coordenadas e_1, e_2, e_3 en tres vectores dados v_1, v_2, v_3? ¿Cuándo es invertible esa matriz?

2.24 Si e_1, e_2, e_3 están en el espacio columna de una matriz de 3 por 5, ¿ésta tiene inversa izquierda? ¿Tiene inversa derecha?

2.25 Suponga que T es una transformación lineal sobre $\mathbf{R}^3$ que transforma cada punto (u, v, w) en $(u + v + w, u + v, u)$. Describa lo que T^{-1} hace al punto (x, y, z).

2.26 *¿Falso o verdadero?*
a) Todo subespacio de $\mathbf{R}^4$ es el espacio nulo de alguna matriz.
b) Si A tiene el mismo espacio nulo que A^{T}, entonces la matriz debe ser cuadrada.
c) La transformación que manda x a $mx + b$ es lineal (de $\mathbf{R}^1$ a $\mathbf{R}^1$).

2.27 Encuentre bases para los cuatro subespacios fundamentales de

$$A_1 = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \quad \text{y} \quad A_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 4 \end{bmatrix}.$$

2.28 *a*) Si los renglones de A son linealmente independientes (A es de m por n), entonces el rango es ____, el espacio columna es ____ y el espacio nulo izquierda es ____.
b) Si A es de 8 por 10 con un espacio nulo bidimensional, demuestre que $Ax = b$ puede resolverse para toda b.

2.29 Describa las transformaciones lineales del plano x-y que están representadas con la base estándar (1, 0) y (0, 1) por las matrices

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

2.30 *a*) Si A es cuadrada, demuestre que el espacio nulo de A^2 contiene al espacio nulo de A.
b) También demuestre que el espacio columna de A^2 está contenido en el espacio columna de A.

2.31 ¿Cuándo se cumple $A^2 = 0$ para la matriz $A = uv^{\mathrm{T}}$ de rango 1?

2.32 *a*) Encuentre una base para el espacio de todos los vectores en $\mathbf{R}^6$ con $x_1 + x_2 = x_3 + x_4 = x_5 + x_6$.
b) Encuentre una matriz cuyo espacio nulo sea ese subespacio.
c) Encuentre una matriz cuyo espacio columna sea ese subespacio.

2.33 Suponga que las matrices $PA = LU$ son

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & -3 & 2 \\ 2 & -1 & 4 & 2 & 1 \\ 4 & -2 & 9 & 1 & 4 \\ 2 & -1 & 5 & -1 & 5 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 4 & 2 & 1 \\ 0 & 0 & 1 & -3 & 2 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

a) ¿Cuál es el rango de A?
b) ¿Cuál es una base para el espacio renglón de A?
c) ¿*Falso o verdadero*?: Los renglones 1, 2, 3 de A son linealmente independientes.
d) ¿Cuál es una base para el espacio columna de A?
e) ¿Cuál es la dimensión del espacio nulo izquierdo de A?
f) ¿Cuál es la solución general de $Ax = 0$?

Capítulo



Ortogonalidad

3.1 VECTORES Y SUBESPACIOS ORTOGONALES

Una base es un conjunto de vectores independientes que genera un espacio. Geométricamente, es un conjunto de ejes coordenados. Un espacio vectorial se define sin estos ejes, aunque cada vez que pienso en el plano x-y, en el espacio tridimensional o en $\mathbf{R}^n$, ahí están los ejes. ¡Suelen ser perpendiculares! *Los ejes coordenados producto de la imaginación prácticamente siempre son ortogonales.* Al elegir una base, se tiende a elegir una ortogonal.

Uno de los fundamentos del álgebra lineal es el concepto de base ortogonal. Se requiere una base para convertir construcciones geométricas en cálculos algebraicos, y se necesita una base ortogonal para que estos cálculos sean sencillos. Especialización más detallada hace casi óptima la base: los vectores deben tener *longitud* 1. Para una ***base ortonormal*** (vectores unitarios ortogonales) se encuentra que

1. la longitud $\|x\|$ de un vector;
2. la prueba $x^{\mathrm{T}}y = 0$ para vectores perpendiculares; y
3. cómo crear vectores perpendiculares a partir de vectores linealmente independientes.

Más que justamente vectores, los *subespacios* también pueden ser perpendiculares. Se descubrirá, de manera tan hermosa y simple que será una delicia ver, que ***los subespacios fundamentales se encuentran a ángulos rectos***. Estos cuatro subespacios fundamentales son perpendiculares por pares, dos en $\mathbf{R}^m$ y dos en $\mathbf{R}^n$. Esto completará el teorema fundamental del álgebra lineal.

El primer paso es encontrar la ***longitud de un vector***, que se denota por $\|x\|$, y en dos dimensiones proviene de la hipotenusa de un triángulo rectángulo (véase la figura 3.1*a*). El cuadrado de la longitud fue proporcionado hace mucho tiempo por Pitágoras: $\|x\|^2 = x_1^2 + x_2^2$.

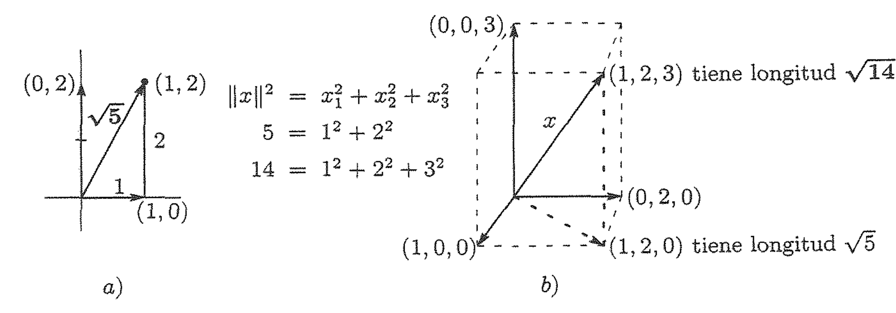


Figura 3.1 Longitud de los vectores (x_1, x_2) y (x_1, x_2, x_3).

En el espacio tridimensional, $x = (x_1, x_2, x_3)$ es la diagonal de una caja (véase la figura 3.1*b*). Su longitud proviene de *dos* aplicaciones de la fórmula de Pitágoras. El caso bidimensional se ocupa de $(x_1, x_2, 0) = (1, 2, 0)$ a través de una base. Esto forma un ángulo recto con el lado vertical $(0, 0, x_3) = (0, 0, 3)$. La hipotenusa del triángulo en negritas (nuevamente Pitágoras) es la longitud $\|x\|$ que se busca:

$$\textbf{Longitud en 3 dimensiones} \quad \|x\|^2 = 1^2 + 2^2 + 3^2 \quad \text{y} \quad \|x\| = \sqrt{x_1^2 + x_2^2 + x_3^2}.$$

La extensión a $x = (x_1, \ldots, x_n)$ en n dimensiones es inmediata. Por el teorema de Pitágoras multiplicado por $n - 1$ veces, ***la longitud de $\|x\|$ en $\mathbf{R}^n$ es la raíz cuadrada positiva de x^Tx:***

$$\textbf{Longitud al cuadrado} \quad \|x\|^2 = x_1^2 + x_2^2 + \cdots + x_n^2 = x^Tx. \tag{1}$$

La suma de cuadrados corresponde a x^Tx, y la longitud de $x = (1, 2, -3)$ es $\sqrt{14}$:

$$x^Tx = \begin{bmatrix} 1 & 2 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} = 1^2 + 2^2 + (-3)^2 = 14.$$

Vectores ortogonales

¿Cómo decidir si dos vectores x y y son perpendiculares? ¿Cuál es la prueba de ortogonalidad en la figura 3.2? En el plano generado por x y y, estos vectores son ortogonales en el supuesto de que formen un *triángulo rectángulo*. Volvamos a $a^2 + b^2 = c^2$:

$$\textbf{Lados de un triángulo rectángulo} \quad \|x\|^2 + \|y\|^2 = \|x - y\|^2. \tag{2}$$

Al aplicar la fórmula de la longitud (1), esta prueba de ortogonalidad en $\mathbf{R}^n$ se vuelve

$$\left(x_1^2 + \cdots + x_n^2\right) + \left(y_1^2 + \cdots + y_n^2\right) = (x_1 - y_1)^2 + \cdots + (x_n - y_n)^2.$$

El miembro derecho contiene un término $-2x_iy_i$ extra de cada $(x_i - y_i)^2$:

$$\text{miembro derecho} = \left(x_1^2 + \cdots + x_n^2\right) - 2(x_1y_1 + \cdots + x_ny_n) + \left(y_1^2 + \cdots + y_n^2\right).$$

Se tiene un triángulo rectángulo cuando la suma de los términos del producto cruz x_iy_i es cero:

$$\textbf{Vectores ortogonales} \quad x^Ty = x_1y_1 + \cdots + x_ny_n = 0. \tag{3}$$

Esta sumatoria es $x^Ty = \Sigma\, x_iy_i = y^Tx$, el vector renglón x^T multiplicado por el vector columna y:

$$\textbf{Producto interno} \quad x^Ty = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = x_1y_1 + \cdots + x_ny_n. \tag{4}$$

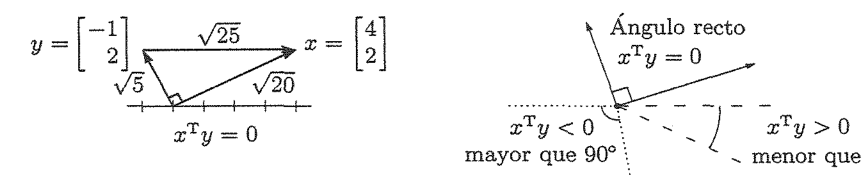


Figura 3.2 Triángulo rectángulo con $5 + 20 = 25$. Ángulo con línea punteada 100°, ángulo con línea discontinua 30°.

Algunas veces este número se denomina producto escalar o producto punto, y se denota por (x, y) o $x \cdot y$. Aquí se usará el nombre de ***producto interno*** y se preservará la notación x^Ty.

3A El producto interno x^Ty es cero si y sólo si x y y son vectores ortogonales. Si $x^Ty > 0$, su ángulo es menor que 90°. Si $x^Ty < 0$, su ángulo es mayor que 90°.

La longitud al cuadrado es el producto interno de x consigo mismo: $x^Tx = x_1^2 + \cdots + x_n^2 = \|x\|^2$. El único vector con longitud cero —el único vector que es ortogonal a sí mismo— es el vector cero. Este vector $x = 0$ es ortogonal a cada vector en $\mathbf{R}^n$.

Ejemplo 1 $(2, 2, -1)$ es ortogonal a $(-1, 2, 2)$. La longitud de cada uno es $\sqrt{4 + 4 + 1} = 3$.

Hecho útil: **Si los vectores diferentes de cero $v_1, \ldots, v_k$ son mutuamente ortogonales** (cada vector es perpendicular a los demás), **entonces estos vectores son linealmente independientes.**

Demostración Suponga que $c_1v_1 + \cdots + c_kv_k = 0$. Para demostrar que c_1 debe ser cero, se toma el producto interno de ambos miembros con v_1. Debido a la ortogonalidad de las vs, se queda con un solo término:

$$v_1^T(c_1v_1 + \cdots + c_kv_k) = c_1v_1^Tv_1 = 0. \tag{5}$$

Los vectores son diferentes de cero, de modo que $v_1^Tv_1 \neq 0$ y en consecuencia $c_1 = 0$. Lo mismo se cumple para cada c_i. La única combinación de las vs que produce 0 tiene a todas las $c_i = 0$: ¡*independencia*! ■

Los vectores coordenados $e_1, \ldots, e_n$ en $\mathbf{R}^n$ son los vectores ortogonales más importantes. Éstos son las columnas de la matriz identidad. Constituyen la base más sencilla de $\mathbf{R}^n$, y son *vectores unitarios*: la longitud de cada uno es $\|e_i\| = 1$. Apuntan a lo largo de los ejes de coordenadas. Si estos vectores se rotan, el resultado es una nueva **base ortonormal**: un nuevo sistema de *vectores unitarios mutuamente ortogonales*. En $\mathbf{R}^2$ se tiene $\cos^2\theta + \operatorname{sen}^2\theta = 1$:

$$\textbf{Vectores ortonormales en } \mathbf{R}^2 \quad v_1 = (\cos\theta, \operatorname{sen}\theta) \text{ y } v_2 = (-\operatorname{sen}\theta, \cos\theta).$$

Subespacios ortogonales

A continuación se abordará la ortogonalidad de dos subespacios. **Cada vector** *en un subespacio debe ser ortogonal a* **cada vector** *en el otro subespacio*. Los subespacios de $\mathbf{R}^3$ pueden tener dimensión 0, 1, 2, o 3. Estos subespacios están representados por rectas o planos que pasan por el origen, y en casos extremos, sólo por el origen o todo el espacio. El subespacio $\{0\}$ es ortogonal a todos los subespacios. Una recta puede ser ortogonal a otra recta, o puede ser ortogonal a un plano, pero *un plano no puede ser ortogonal a un plano*.

Debo admitir que el muro frontal y el muro lateral de una habitación se ven como planos perpendiculares en $\mathbf{R}^3$. Pero según nuestra definición, ¡no es así! Hay rectas v y w en los muros frontal y lateral que no se encuentran a ángulos rectos. La recta que corre a lo largo de la arista pertenece a *ambos* muros, y ciertamente no es ortogonal a sí misma.

3B Dos subespacios $\mathbf{V}$ y $\mathbf{W}$ del mismo espacio $\mathbf{R}^n$ *son ortogonales* si cada vector v en $\mathbf{V}$ es ortogonal a cada vector w en $\mathbf{W}$: $v^{\mathrm{T}}w = 0$ para toda v y w.

Ejemplo 2 Suponga que $\mathbf{V}$ es el plano generado por $v_1 = (1, 0, 0, 0)$ y $v_2 = (1, 1, 0, 0)$. Si $\mathbf{W}$ es el plano generado por $w = (0, 0, 4, 5)$, entonces w es ortogonal a ambos vectores vs. La recta $\mathbf{W}$ es ortogonal a todo el plano $\mathbf{V}$.

En este caso, con subespacios de dimensiones 2 y 1 en $\mathbf{R}^4$, hay espacio para un tercer subespacio. La recta $\mathbf{L}$ que pasa por $z = (0, 0, 5, -4)$ es perpendicular a $\mathbf{V}$ y $\mathbf{W}$. Así, la suma de las dimensiones es $2 + 1 + 1 = 4$. ¿Qué espacio es perpendicular a los tres, $\mathbf{V}$, $\mathbf{W}$, y $\mathbf{L}$?

Los importantes subespacios ortogonales no se presentan accidentalmente, y se presentan por pares. De hecho, los subespacios ortogonales son inevitables: ***¡Son los espacios fundamentales!*** El primer par lo constituyen el *espacio nulo* y el *espacio renglón*. Estos son subespacios de $\mathbf{R}^n$: los renglones tienen n componentes, como el vector x en $Ax = 0$. Es necesario demostrar, usando $Ax = 0$, que ***los renglones de A son ortogonales al vector x en el espacio nulo.***

3C Teorema fundamental de ortogonalidad El espacio renglón es ortogonal al espacio nulo (en $\mathbf{R}^n$). El espacio columna es ortogonal al espacio nulo izquierdo (en $\mathbf{R}^m$).

Primera demostración Suponga que x es un vector en el espacio nulo. Entonces $Ax = 0$, y este sistema de m ecuaciones puede escribirse como renglones de A que multiplican a x:

$$\textbf{Todo renglón es ortogonal a } x \qquad Ax = \begin{bmatrix} \cdots & \textbf{renglón 1} & \cdots \\ \cdots & \text{renglón 2} & \cdots \\ & \vdots & \\ \cdots & \text{renglón } m & \cdots \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{6}$$

La cuestión importante ya está presente en la primera ecuación: *el renglón 1 es ortogonal a x*. Su producto interno es cero: esa es la ecuación 1. Todo miembro derecho es cero, de modo que x es ortogonal a todo renglón. En consecuencia, x es ortogonal a *toda combinación* de los renglones. Cada x en el espacio nulo es ortogonal a cada vector en el espacio renglón, de modo que $N(A) \perp C(A^{\mathrm{T}})$.

El otro par de subespacios ortogonales proviene de $A^{\mathrm{T}}y = 0$, o $y^{\mathrm{T}}A = 0$:

$$y^{\mathrm{T}}A = \begin{bmatrix} y_1 & \cdots & y_m \end{bmatrix} \begin{bmatrix} \mathbf{c} & & \mathbf{c} \\ \mathbf{o} & & \mathbf{o} \\ \mathbf{l} & & \mathbf{l} \\ \mathbf{u} & \cdots & \mathbf{u} \\ \mathbf{m} & & \mathbf{m} \\ \mathbf{n} & & \mathbf{n} \\ \mathbf{a} & & \mathbf{a} \\ \mathbf{1} & & n \end{bmatrix} = \begin{bmatrix} 0 & \cdots & 0 \end{bmatrix}. \tag{7}$$

El vector y es ortogonal a cada columna. La ecuación así lo indica, a partir de los ceros en el miembro derecho. En consecuencia, y es ortogonal a toda combinación de las columnas.

Es ortogonal al espacio columna, y se trata de un vector típico en el espacio nulo izquierdo: $N(A^{\mathrm{T}}) \perp C(A)$. Esto es lo mismo que la primera mitad del teorema, con A sustituida por A^{T}. ■

Segunda demostración El contraste con esta "demostración libre de coordenadas" debe ser de utilidad para el lector. Constituye un método de razonamiento más "abstracto". Me gustaría saber cuál de las dos demostraciones es más clara y mejor comprendida.

Si x está en el espacio nulo, entonces $Ax = 0$. Si v está en el espacio renglón, es una combinación de los renglones: $v = A^{\mathrm{T}}z$ para algún vector z. Así, en una línea:

$$\textbf{Espacio nulo} \perp \textbf{Espacio renglón} \qquad v^{\mathrm{T}}x = (A^{\mathrm{T}}z)^{\mathrm{T}}x = z^{\mathrm{T}}Ax = z^{\mathrm{T}}0 = 0. \tag{8}$$

■

Ejemplo 3 Suponga que el rango de A es 1, de modo que su espacio columna y su espacio renglón son rectas:

$$\textbf{Matriz de rango 1} \qquad A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \\ 3 & 9 \end{bmatrix}.$$

Los renglones son múltiplos de $(1, 3)$. El espacio nulo contiene a $x = (-3, 1)$, que es ortogonal a todos los renglones. El espacio nulo y el espacio renglón son rectas perpendiculares en $\mathbf{R}^2$:

$$\begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = 0 \quad \text{y} \quad \begin{bmatrix} 2 & 6 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = 0 \quad \text{y} \quad \begin{bmatrix} 3 & 9 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = 0.$$

En contraste, los otros dos subespacios están en $\mathbf{R}^3$. El espacio columna es la recta que pasa por $(1, 2, 3)$. El espacio nulo izquierdo debe ser el *plano* $y_1 + 2y_2 + 3y_3 = 0$. Esta ecuación es exactamente el contenido de $y^{\mathrm{T}}A = 0$.

Los dos primeros subespacios (las dos rectas) tienen dimensiones $1 + 1 = 2$ en el espacio $\mathbf{R}^2$. El segundo par (recta y plano) tienen dimensiones $1 + 2 = 3$ en el espacio $\mathbf{R}^3$. En general, *el espacio renglón y el espacio nulo tienen dimensiones cuya suma es* $r + (n - r) = n$. La suma de las dimensiones del otro par es igual a $r + (m - r) = m$. Está ocurriendo algo más que la ortogonalidad, por lo que pido paciencia al lector para abordar otra cuestión adicional: **las dimensiones**.

Con toda certeza es verdad que el espacio nulo es perpendicular al espacio renglón, aunque ésta no es toda la verdad. $N(A)$ *contiene a todo vector ortogonal al espacio renglón.* El espacio nulo fue formado a partir de *todas* las soluciones de $Ax = 0$.

DEFINICIÓN Dado un subespacio $\mathbf{V}$ de $\mathbf{R}^n$, el espacio de *todos* los vectores ortogonales a $\mathbf{V}$ se denomina **complemento ortogonal** de $\mathbf{V}$. Se denota por $\mathbf{V}^{\perp}$ = "$\mathbf{V}$ perp."

Con esta terminología, el espacio nulo es el complemento ortogonal del espacio renglón: $N(A) = (C(A^{\mathrm{T}}))^{\perp}$. Al mismo tiempo, el espacio renglón contiene a todos los vectores que son ortogonales al espacio nulo. Un vector z no puede ser ortogonal al espacio nulo sino fuera del espacio renglón. Al agregar z como un renglón extra de A se agranda el espacio renglón, aunque se sabe que hay una fórmula fija $r + (n - r) = n$:

Fórmula de la dimensión dim(espacio renglón) + dim(espacio nulo) = número de columnas.

Todo vector ortogonal al espacio nulo está en el espacio renglón: $C(A^{\mathrm{T}}) = (N(A))^{\perp}$.

El mismo razonamiento aplicado a A^T produce el resultado dual: *El espacio nulo izquierdo $N(A^T)$ y el espacio columna $C(A)$ son complementos ortogonales.* La suma de sus dimensiones es igual a $(m - r) + r = m$. Esto completa la segunda mitad del teorema fundamental del álgebra lineal. La primera mitad proporcionó las dimensiones de los cuatro subespacios, incluyendo el hecho de que rango renglón = rango columna. Ahora se sabe que estos subespacios son perpendiculares. Más que eso, los subespacios son complementos ortogonales.

3D Teorema fundamental del álgebra lineal, parte II

El espacio nulo es el *complemento ortogonal* del espacio renglón en $\mathbf{R}^n$.
El espacio nulo izquierdo es el *complemento ortogonal* del espacio columna en $\mathbf{R}^m$.

Para repetir, el espacio renglón contiene todo lo que es ortogonal al espacio nulo. El espacio columna contiene todo lo que es ortogonal al espacio nulo izquierdo. Esta es sólo una oración, escondida a medio libro, pero ¡*decide cuáles ecuaciones pueden resolverse*! Si se mira directamente, $Ax = b$ requiere que b esté en el espacio columna. Si se mira indirectamente, $Ax = b$ **requiere que b sea perpendicular al espacio nulo izquierdo**.

3E $Ax = b$ es resoluble si y sólo si $y^Tb = 0$ siempre que $y^TA = 0$.

El método directo fue "b debe ser una combinación de las columnas". El método indirecto es "b debe ser ortogonal a cada vector ortogonal a las columnas". Esto no se ve como una mejoría (para plantearlo en términos suaves). Pero si sólo uno o dos vectores son ortogonales a las columnas, resulta mucho más fácil comprobar estas una o dos condiciones $y^Tb = 0$. Un buen ejemplo es la ley del voltaje de Kirchhoff en la sección 2.5. Verificar cero alrededor de los circuitos es mucho más fácil que reconocer combinaciones de las columnas.

Cuando la suma de los miembros izquierdos de $Ax = b$ es cero, entonces los miembros derechos también tienen que cumplir lo siguiente:

$$\begin{aligned} x_1 - x_2 &= b_1 \\ x_2 - x_3 &= b_2 \\ x_3 - x_1 &= b_3 \end{aligned} \quad \text{es resoluble si y sólo si } b_1 + b_2 + b_3 = 0. \text{ Aquí } A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}.$$

Esta prueba $b_1 + b_2 + b_3 = 0$ hace que b sea ortogonal a $y = (1, 1, 1)$ en el espacio nulo izquierdo. Por el teorema fundamental, ¡*b es una combinación de las columnas*!

La matriz y los subespacios

Se recalca que $\mathbf{V}$ y $\mathbf{W}$ pueden ser ortogonales sin ser complementos. Sus dimensiones también pueden ser pequeñas. La recta $\mathbf{V}$ generada por $(0, 1, 0)$ es ortogonal a la recta $\mathbf{W}$ generada por $(0, 0, 1)$, pero $\mathbf{V}$ no es $\mathbf{W}^\perp$. El complemento ortogonal de $\mathbf{W}$ es un plano bidimensional, y la recta es sólo parte de $\mathbf{W}^\perp$. Cuando las dimensiones están correctas, los subespacios ortogonales *son* necesariamente complementos ortogonales.

$$\text{Si } \mathbf{W} = \mathbf{V}^\perp, \text{ entonces } \mathbf{V} = \mathbf{W}^\perp \text{ y } \dim \mathbf{V} + \dim \mathbf{W} = n.$$

En otras palabras, $\mathbf{V}^{\perp\perp} = \mathbf{V}$. Las dimensiones de $\mathbf{V}$ y $\mathbf{W}$ están correctas, y todo el espacio $\mathbf{R}^n$ se ha descompuesto en dos partes perpendiculares (véase la figura 3.3).

W — Dos ejes ortogonales en $\mathbf{R}^3$ — No son complementos ortogonales — V

W — Recta $\mathbf{W}$ perpendicular al plano $\mathbf{V}$ — Complementos ortogonales $\mathbf{V} = \mathbf{W}^\perp$ — V

Figura 3.3 Complementos ortogonales en $\mathbf{R}^3$: un plano y una recta (no dos rectas).

Al separar $\mathbf{R}^n$ en dos partes ortogonales, cada vector se separa en $x = v + w$. El vector v es la proyección sobre el subespacio $\mathbf{V}$. La componente ortogonal w es la proyección de x sobre $\mathbf{W}$. En la siguiente sección se muestra cómo encontrar estas proyecciones de x. Éstas conducen a lo que quizá es la figura más importante del libro (véase la figura 3.4).

En la figura 3.4 se resume el teorema fundamental del álgebra lineal. Ilustra el verdadero efecto de una matriz: lo que ocurre al interior de la multiplicación Ax. El espacio nulo es llevado al vector cero. Todo Ax está en el espacio columna. Nada se manda al espacio nulo izquierdo. *La verdadera acción es entre el espacio renglón y el espacio columna*, y el lector verá esto al analizar un vector típico x. Éste contiene una " componente del espacio renglón" y una "componente del espacio nulo", con $x = x_r + x_n$. Cuando esto se multiplica por A, se trata de $Ax = Ax_r + Ax_n$:

La componente del espacio nulo se va en cero: $Ax_n = 0$.
La componente del espacio renglón se va al espacio columna: $Ax_r = Ax$.

Por supuesto, todo va al espacio columna: la matriz no puede hacer otra cosa. He intentado hacer los espacios renglón y columna del mismo tamaño, con dimensión igual r.

3F Del espacio renglón al espacio columna, A en realidad es invertible. Cada vector b en el espacio columna proviene exactamente de un vector x_r en el espacio renglón.

dim r — espacio renglón de A — x_r — $Ax_r = b$ — espacio columna de A — b — dim r — $x = x_r + x_n$ — $Ax = b$ — $\mathbf{R}^n$ — $\mathbf{R}^m$ — x_n — $Ax_n = 0$ — espacio nulo — dim $n - r$ — espacio nulo izquierdo — dim $m - r$

Figura 3.4 La verdadera acción $Ax = A(x_{\text{renglón}} + x_{\text{nulo}})$ de cualquier matriz de m por n.

Demostración Todo b en el espacio columna es una combinación Ax de las columnas. De hecho, b es Ax_r, con x_r en el espacio renglón, ya que la componente del espacio nulo da $Ax_n = 0$. Si otro vector x'_r en el espacio renglón proporciona $Ax'_r = b$, entonces $A(x_r - x'_r) = b - b = 0$. Esto coloca a $x_r - x'_r$ en el espacio nulo y en el espacio renglón, lo cual lo hace ortogonal a sí mismo. En consecuencia, es cero, y $x_r = x'_r$. Exactamente un vector en el espacio renglón es llevado a b. ■

Toda matriz transforma su espacio renglón sobre su espacio columna.

Sobre estos espacios r-dimensionales, A es invertible. Sobre su espacio nulo, A es cero. Cuando A es diagonal, se observa que la submatriz invertible preserva los r no ceros.

A^T va en dirección opuesta, de $\mathbf{R}^m$ a $\mathbf{R}^n$ y de $C(A)$ de vuelta a $C(A^T)$. Por supuesto, ¡la traspuesta no es la inversa! A^T mueve correctamente los espacios, pero no los vectores individuales. Este honor corresponde a A^{-1} en caso de existir, y sólo existe si $r = m = n$. No puede pedirse que A^{-1} regrese todo un espacio nulo a partir del vector cero.

Cuando A^{-1} no existe, el mejor sustituto es la ***seudoinversa*** A^+. Esta matriz invierte A donde es posible hacerlo: $A^+Ax = x$ para x en el espacio renglón. Sobre el espacio nulo izquierdo, nada puede hacerse: $A^+y = 0$. Así, A^+ invierte A cuando ésta es invertible, y tiene el mismo rango r. Una fórmula para A^+ depende de la ***descomposición del valor singular***, tema para el cual antes se requiere saber algo sobre valores característicos (o singulares).

Conjunto de problemas 3.1

1. Encuentre las longitudes y el producto interno de $x = (1, 4, 0, 2)$ y $y = (2, -2, 1, 3)$.
2. Proporcione un ejemplo en $\mathbf{R}^2$ de vectores linealmente independientes que no sean ortogonales. También, proporcione un ejemplo de vectores ortogonales que no sean independientes.
3. Dos rectas en el plano son perpendiculares cuando el producto de sus pendientes es -1. Aplique este hecho a los vectores $x = (x_1, x_2)$ y $y = (y_1, y_2)$, cuyas pendientes son x_2/x_1 y y_2/y_1, para deducir nuevamente la condición de ortogonalidad $x^Ty = 0$.
4. Cómo puede saberse si el i-ésimo renglón de una matriz invertible B es ortogonal a la j-ésima columna de B^{-1}, si $i \neq j$?
5. De los siguientes vectores, v_1, v_2, v_3, v_4 ¿qué pares son ortogonales?

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ -2 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 4 \\ 0 \\ 4 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

6. Encuentre todos los vectores en $\mathbf{R}^3$ que son ortogonales a $(1, 1, 1)$ y $(1, -1, 0)$. Escriba una base ortonormal a partir de estos vectores (vectores unitarios mutuamente ortogonales).
7. Encuentre un vector x ortogonal al espacio renglón de A, un vector y ortogonal al espacio columna, y un vector z ortogonal al espacio nulo:

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 3 \\ 3 & 6 & 4 \end{bmatrix}.$$

8. Si $\mathbf{V}$ y $\mathbf{W}$ son subespacios ortogonales, demuestre que el único vector que tienen en común es el vector cero: $\mathbf{V} \cap \mathbf{W} = \{0\}$.
9. Encuentre el complemento ortogonal del plano generado por los vectores $(1, 1, 2)$ y $(1, 2, 3)$, considerando que son los renglones de A y resolviendo $Ax = 0$. Recuerde que el complemento es toda una recta.
10. Escriba una ecuación homogénea en tres incógnitas cuyas soluciones sean la combinación lineal de los vectores $(1, 1, 2)$ y $(1, 2, 3)$. Ésta es el inverso del ejercicio previo, aunque en realidad ambos problemas son el mismo.
11. A menudo, el teorema fundamental se plantea en forma de la *alternativa de Fredholm*: Para A y b cualesquiera, uno y sólo uno de los siguientes sistemas tiene una solución:

    i) $Ax = b$.

    ii) $A^Ty = 0$, $y^Tb \neq 0$.

    Ya sea que b esté en el espacio columna $C(A)$ o que en $N(A^T)$ exista una y tal que $y^Tb \neq 0$. Demuestre que esto es contradictorio, ya que ambos i) y ii) tienen soluciones.
12. Encuentre una base para el complemento ortogonal del espacio renglón de A:

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 4 \end{bmatrix}.$$

    Separe $x = (3, 3, 3)$ en una componente del espacio renglón x_r y una componente del espacio nulo x_n.
13. Ilustre la acción de A^T mediante una figura que corresponda a la figura 3.4, mandando $C(A)$ de vuelta al espacio renglón y el espacio nulo izquierdo a cero.
14. Demuestre que $x - y$ es ortogonal a $x + y$ si y sólo si $\|x\| = \|y\|$.
15. Encuentre una matriz cuyo espacio renglón contenga a $(1, 2, 1)$ y cuyo espacio nulo a $(1, -2, 1)$, o demuestre que no puede existir una matriz así.
16. Encuentre todos los vectores que son perpendiculares a $(1, 4, 4, 1)$ y $(2, 9, 8, 2)$.
17. Si $\mathbf{V}$ es el complemento ortogonal de $\mathbf{W}$ en $\mathbf{R}^n$, ¿existe una matriz con espacio renglón $\mathbf{V}$ y espacio nulo $\mathbf{W}$? Empiece con una base de $\mathbf{V}$, y obtenga tal matriz.
18. Si $\mathbf{S} = \{0\}$ es el subespacio de $\mathbf{R}^4$ que sólo contiene al vector cero, ¿cuál es $\mathbf{S}^\perp$? Si $\mathbf{S}$ es generado por $(0, 0, 0, 1)$, ¿cuál es $\mathbf{S}^\perp$? ¿Cuál es $(\mathbf{S}^\perp)^\perp$?
19. *¿Por qué son falsas las siguientes afirmaciones?*

    a) Si $\mathbf{V}$ es ortogonal a $\mathbf{W}$, entonces $\mathbf{V}^\perp$ es ortogonal a $\mathbf{W}^\perp$.

    b) $\mathbf{V}$ es ortogonal a $\mathbf{W}$ y $\mathbf{W}$ ortogonal a $\mathbf{Z}$ hace que $\mathbf{V}$ sea ortogonal a $\mathbf{Z}$.
20. Sea $\mathbf{S}$ un subespacio de $\mathbf{R}^n$. Explique el significado de $(\mathbf{S}^\perp)^\perp = \mathbf{S}$, y por qué es verdadera.
21. Sea $\mathbf{P}$ el plano en $\mathbf{R}^2$ con ecuación $x + 2y - z = 0$. Encuentre un vector perpendicular a $\mathbf{P}$. ¿Qué matriz tiene al plano $\mathbf{P}$ como espacio nulo y qué matriz tiene a $\mathbf{P}$ como su espacio renglón?
22. Sea $\mathbf{S}$ el plano de $\mathbf{R}^4$ que contiene a todos los vectores con $x_1 + x_2 + x_3 + x_4 = 0$. Encuentre una base para el espacio $\mathbf{S}^\perp$ que contenga a todos los vectores ortogonales a $\mathbf{S}$.

23. Construya una matriz no simétrica de 2 por 2 con rango 1. Copie la figura 3.4 y escriba un vector en cada subespacio. ¿Qué vectores son ortogonales?

24. Vuelva a dibujar la figura 3.4 para una matriz de 3 por 2 con rango $r = 2$. ¿Cuál subespacio es $\boldsymbol{Z}$ (sólo el vector cero)? La parte del espacio nulo de cualquier vector x en $\mathbf{R}^2$ es $x_n =$ ____.

25. Construya una matriz con la propiedad requerida o escriba una razón que explique por qué es imposible hacerlo.

a) El espacio columna contiene a $\begin{bmatrix}1\\2\\-3\end{bmatrix}$ y $\begin{bmatrix}2\\-3\\5\end{bmatrix}$, el espacio columna contiene a $\begin{bmatrix}1\\1\\1\end{bmatrix}$.

b) El espacio renglón contiene a $\begin{bmatrix}1\\2\\-3\end{bmatrix}$ y $\begin{bmatrix}2\\-3\\5\end{bmatrix}$, el espacio nulo contiene a $\begin{bmatrix}1\\1\\1\end{bmatrix}$.

c) $Ax = \begin{bmatrix}1\\1\\1\end{bmatrix}$ tiene una solución y $A^{\mathrm{T}}\begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$.

d) Cada renglón es ortogonal a cada columna (A no es la matriz cero).

e) La suma de las columnas es una columna de 0s y la suma de los renglones es un renglón de 1s.

26. Si $AB = 0$, entonces las columnas de B están en las (los)____ de A. Los renglones de A están en las (los) ____ de B. ¿Por qué A y B no pueden ser matrices de 3 por 3 con rango 2?

27. *a)* Si $Ax = b$ tiene una solución y $A^{\mathrm{T}}y = 0$, entonces y es perpendicular a ____.
b) Si $A^{\mathrm{T}}y = c$ tiene una solución y $Ax = 0$, entonces x es perpendicular a ____.

28. El siguiente es un sistema de ecuaciones $Ax = b$ *sin solución*:

$$\begin{aligned} x + 2y + 2z &= 5 \\ 2x + 2y + 3z &= 5 \\ 3x + 4y + 5z &= 9. \end{aligned}$$

Encuentre números y_1, y_2 y y_3 que multipliquen las ecuaciones de modo que sumen $0 = 1$. ¿Encontró un vector y, en cuál subespacio? El producto interno $y^{\mathrm{T}}b$ es 1.

29. En la figura 3.4, ¿cómo se sabe que Ax_r es igual a Ax? ¿Cómo se sabe que este vector está en el espacio columna? Si $A = \begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}$ y $x = \begin{bmatrix}1\\0\end{bmatrix}$, ¿cuál es x_r?

30. Si Ax está en el espacio nulo de A^{T}, entonces $Ax = 0$. Razón: Ax también está en el _____ de A y los espacios son ____. *Conclusión: $A^{\mathrm{T}}A$ tiene el mismo espacio nulo que A.*

31. Suponga que A es una matriz simétrica ($A^{\mathrm{T}} = A$).
a) ¿Por qué su espacio columna es perpendicular a su espacio nulo?
b) Si $Ax = 0$ y $Az = 5z$, ¿cuáles subespacios contienen a los "vectores característicos" x y z? **Las matrices simétricas tienen vectores característicos perpendiculares** (consulte la sección 5.5).

32. (Recomendado) Dibuje la figura 3.4 con la finalidad de mostrar cada subespacio para

$$A = \begin{bmatrix}1 & 2\\3 & 6\end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix}1 & 0\\3 & 0\end{bmatrix}.$$

33. Encuentre las piezas x_r y x_n, y trace idóneamente la figura 3.4, si

$$A = \begin{bmatrix}1 & -1\\0 & 0\\0 & 0\end{bmatrix} \quad \text{y} \quad x = \begin{bmatrix}2\\0\end{bmatrix}.$$

Los problemas 34 a 44 son sobre subespacios ortogonales.

34. Escriba bases de los espacios ortogonales $\mathbf{V}$ y $\mathbf{W}$ en las columnas de las matrices V y W. *¿Por qué $V^{\mathrm{T}}W =$ matriz cero?* Esto corresponde a $v^{\mathrm{T}}w = 0$ para vectores.

35. El piso y la pared no son subespacios ortogonales porque comparten un vector diferente de cero (a lo largo de la recta en que se unen). ¡Dos planos en $\mathbf{R}^3$ no pueden ser ortogonales! Encuentre un vector en ambos espacios columna $C(A)$ y $C(B)$:

$$A = \begin{bmatrix}1 & 2\\1 & 3\\1 & 2\end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix}5 & 4\\6 & 3\\5 & 1\end{bmatrix}.$$

Este es un vector Ax y también $B\hat{x}$. Piense en la matriz $[A\ B]$ de 3 por 4.

36. Extienda el problema 35 a un subespacio p-dimensional $\mathbf{V}$ y un subespacio q-dimensional $\mathbf{W}$ de $\mathbf{R}^n$. ¿Qué desigualdad sobre $p + q$ garantiza que $\mathbf{V}$ se corta con $\mathbf{W}$ en un vector diferente de cero? Estos subespacios no pueden ser ortogonales.

37. Demuestre que cada y en $N(A^{\mathrm{T}})$ es perpendicular a cada Ax en el espacio columna, usando la abreviatura para matrices de la ecuación (8). Empiece con $A^{\mathrm{T}}y = 0$.

38. ¿Si $\mathbf{S}$ es el subespacio de $\mathbf{R}^3$ que sólo contiene al vector cero, ¿cuál es $\mathbf{S}^{\perp}$? Si $\mathbf{S}$ es generado por (1, 1, 1), ¿cuál es $\mathbf{S}^{\perp}$? Si $\mathbf{S}$ es generado por (2, 0, 0) y (0, 0, 3), ¿cuál es $\mathbf{S}^{\perp}$?

39. Suponga que $\mathbf{S}$ sólo contiene a (1, 5, 1) y a (2, 2, 2) (no es un subespacio). Entonces $\mathbf{S}^{\perp}$ es el espacio nulo de la matriz $A =$ ____. $\mathbf{S}^{\perp}$ es un subespacio aun cuando $\mathbf{S}$ no lo sea.

40. Suponga que $\mathbf{L}$ es un subespacio unidimensional (una recta) en $\mathbf{R}^3$. Su complemento ortogonal $\mathbf{L}^{\perp}$ es el (la)____ perpendicular a $\mathbf{L}$. Luego, $(\mathbf{L}^{\perp})^{\perp}$ es un (una) ____ perpendicular a $\mathbf{L}^{\perp}$. De hecho, $(\mathbf{L}^{\perp})^{\perp}$ es lo mismo que ____.

41. Suponga que $\mathbf{V}$ es todo el espacio $\mathbf{R}^4$. Entonces $\mathbf{V}^{\perp}$ sólo contiene al vector ____. Así, $(\mathbf{V}^{\perp})^{\perp}$ es ____. De modo que $(\mathbf{V}^{\perp})^{\perp}$ es lo mismo que ____.

42. Suponga que $\mathbf{S}$ es generado por los vectores (1, 2, 2, 3) y (1, 3, 3, 2). Encuentre dos vectores que generen a $\mathbf{S}^{\perp}$. Esto es lo mismo que resolver $Ax = 0$ ¿para cuál A?

43. Si $\mathbf{P}$ es el plano de los vectores en $\mathbf{R}^4$ que satisfacen $x_1 + x_2 + x_3 + x_4 = 0$, escriba una base de $\mathbf{P}^{\perp}$. Escriba una matriz cuyo espacio nulo sea $\mathbf{P}$.

44. Si un subespacio $\mathbf{S}$ está contenido en un subespacio $\mathbf{V}$, demuestre que $\mathbf{S}^{\perp}$ contiene a $\mathbf{V}^{\perp}$.

Los problemas 45 a 50 son sobre columnas y renglones perpendiculares.

45. Suponga que una matriz de n por n es invertible: $AA^{-1} = I$. Entonces, la primera columna de A^{-1} es ortogonal al espacio generado por ¿cuáles renglones de A?

46. Encuentre $A^{\mathrm{T}}A$ si las columnas de A son vectores unitarios, todos mutuamente perpendiculares.

47. Construya una matriz A de 3 por 3 sin elementos cero cuyas columnas sean mutuamente perpendiculares. Calcule $A^{\mathrm{T}}A$. ¿Por qué es una matriz diagonal?

48. Las rectas $3x + y = b_1$ y $6x + 2y = b_2$ son ____. Son la misma recta si ____. En ese caso, (b_1, b_2) es perpendicular al vector ____. El espacio nulo de la matriz es la recta $3x + y =$ ____. Un vector particular en ese espacio nulo es ____.

49. ¿Por qué es falsa cada una de las siguientes afirmaciones?
a) (1, 1, 1) es perpendicular a (1, 1, −2), de modo que los planos $x + y + z = 0$ y $x + y - 2z = 0$ son subespacios ortogonales.
b) El subespacio generado por (1, 1, 0, 0, 0) y (0, 0, 0, 1, 1) es el complemento ortogonal del subespacio generado por (1, −1, 0, 0, 0) y (2, −2, 3, 4, −4).
c) Dos subespacios que sólo se cortan en cero son ortogonales.

50. Encuentre una matriz con $v = (1, 2, 3)$ en el espacio renglón y en el espacio columna. Encuentre otra matriz con v en el espacio nulo y en el espacio columna. ¿En cuáles pares de subespacios puede *no* estar v?

51. Suponga que A es de 3 por 4, que B es de 4 por 5, y que $AB = 0$. Demuestre que rango(A) + rango(B) ≤ 4.

52. El comando N = null(A) produce una base del espacio nulo de A. Así, el comando B = null(N') produce una base para el ____ de A.

3.2 COSENOS Y PROYECCIONES SOBRE RECTAS

Los vectores con $x^{\mathrm{T}}y = 0$ son ortogonales. Ahora se quiere permitir productos internos ***distintos de cero***, y ángulos que ***no sean rectos***. Se desea relacionar los productos internos con los ángulos, y también con las traspuestas. En el capítulo 1 la traspuesta se construyó dándole la vuelta a una matriz como si fuese algún tipo de panqué. Es necesario mejorar lo anterior.

Algo es inevitable: *el caso ortogonal es el más importante*. Suponga que se desea encontrar la distancia de un punto b a la recta en dirección del vector a. En esa recta se busca el punto p más próximo a b. La clave está en la geometría: ***la recta que une b con p*** (la recta discontinua en la figura 3.5) ***es perpendicular* a *a*.** *Este hecho permite encontrar la proyección p.* Aun cuando a y b no sean ortogonales, el problema de la distancia automáticamente supone ortogonalidad.

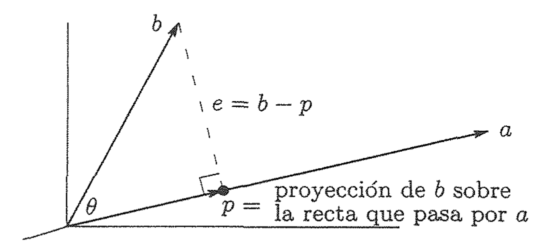


Figura 3.5 La proyección p es el punto (sobre la recta que pasa por a) más próximo a b.

La situación es la misma que cuando se tiene un plano (o con cualquier subespacio S) en vez de una recta. De nuevo, el problema es encontrar el punto p en ese subespacio que sea el más próximo a b. ***Este punto p es la proyección de b sobre el subespacio.*** Una recta perpendicular de b a S corta al subespacio en p. Geométricamente, así se obtiene la distancia entre puntos b y subespacios S. Sin embargo, hay dos preguntas que es necesario plantear:

1. Esta proyección, ¿realmente se presenta en aplicaciones prácticas?
2. Si se tiene una base del subespacio S, ¿hay una fórmula para la proyección p?

Ciertamente, las respuestas son sí. Este es exactamente el problema de la ***solución por mínimos cuadrados de un sistema sobredeterminado***. El vector b representa los datos provenientes de experimentos o cuestionarios, y contiene demasiados errores para ser encontrado en el subespacio S. Cuando se intenta escribir b como una combinación de los vectores de la base de S, no es posible hacerlo, ya que las ecuaciones son inconsistentes, y $Ax = b$ no tiene solución.

El método de mínimos cuadrados selecciona a p como la mejor opción para sustituir a b. No puede haber duda sobre la importancia de esta aplicación. En economía y estadística, los mínimos cuadrados introducen el *análisis de regresión*. En geodesia, el departamento de agrimensura de Estados Unidos abordó 2.5 millones de ecuaciones en 400 000 incógnitas.

Una fórmula para p es fácil cuando el subespacio es una recta. Entonces b se proyectará sobre a de varias formas, y la proyección p se relacionará con varios productos internos y ángulos. La proyección sobre un subespacio de dimensión mayor es por mucho el caso más importante; corresponde a un problema de mínimos cuadrados con varios parámetros, y se resolverá en la sección 3.3. Las fórmulas son incluso más sencillas cuando se produce una base ortogonal de S.

Productos internos y cosenos

A continuación se abordará el análisis de los productos internos y los ángulos. Pronto se verá que no es el ángulo, sino ***el coseno del ángulo***, el que está relacionado directamente con los productos internos. Se acude a la trigonometría en el caso bidimensional para encontrar la relación. Suponga que los vectores a y b forman ángulos α y β con el eje x (véase la figura 3.6).

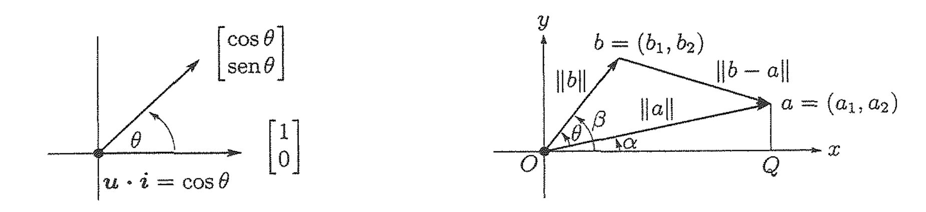


Figura 3.6 Coseno del ángulo $\theta = \beta - \alpha$ usando productos internos.

La longitud $\|a\|$ es la hipotenusa del triángulo OaQ. Así, el seno y el coseno de α son

$$\operatorname{sen}\alpha = \frac{a_2}{\|a\|}, \qquad \cos\alpha = \frac{a_1}{\|a\|}.$$

Para el ángulo β, el seno es $b_2/\|b\|$, y el coseno es $b_1/\|b\|$. El coseno de $\theta = \beta - \alpha$ proviene de una identidad inolvidable:

$$\textbf{Fórmula del coseno} \qquad \cos\theta = \cos\beta\cos\alpha + \operatorname{sen}\beta\operatorname{sen}\alpha = \frac{a_1b_1 + a_2b_2}{\|a\|\,\|b\|}. \tag{1}$$

El numerador en esta fórmula es exactamente el producto interno de a y b. Esto es proporcional a la relación entre $a^{\mathrm{T}}b$ y $\cos\theta$:

3G El coseno del ángulo entre dos vectores a y b cualesquiera diferentes de cero es

Coseno de θ $$\cos\theta = \frac{a^{\mathrm{T}}b}{\|a\|\,\|b\|}. \quad (2)$$

Esta fórmula es dimensionalmente correcta: si se duplica la longitud de b, entonces tanto el numerador como el denominador se duplican, y el coseno permanece sin cambio. Al invertir el signo de b, por otra parte, se invierte el signo de $\cos\theta$ y el ángulo cambia por 180°.

Hay otra ley de trigonometría que conduce directamente al mismo resultado. No es tan inolvidable como la fórmula de la ecuación (1), aunque relaciona las longitudes de los lados de cualquier triángulo:

Ley de los cosenos $$\|b-a\|^2 = \|b\|^2 + \|a\|^2 - 2\|b\|\,\|a\|\cos\theta. \quad (3)$$

Cuando θ es un ángulo recto, se acude al teorema de Pitágoras: $\|b-a\|^2 = \|b\|^2 + \|a\|^2$. Para cualquier ángulo θ, la expresión $\|b-a\|^2$ es $(b-a)^{\mathrm{T}}(b-a)$, y la ecuación (3) se convierte en

$$b^{\mathrm{T}}b - 2a^{\mathrm{T}}b + a^{\mathrm{T}}a = b^{\mathrm{T}}b + a^{\mathrm{T}}a - 2\|b\|\,\|a\|\cos\theta.$$

Al cancelar $b^{\mathrm{T}}b$ y $a^{\mathrm{T}}a$ en ambos miembros de esta ecuación, se reconoce la fórmula (2) para el coseno: $a^{\mathrm{T}}b = \|a\|\,\|b\|\cos\theta$. De hecho, esto demuestra la fórmula del coseno en n dimensiones, ya que sólo es necesario preocuparse sobre el triángulo plano Oab.

Proyección sobre una recta

Ahora se desea encontrar el punto proyección p. Este punto debe ser algún múltiplo $p = \hat{x}a$ del vector dado a: cada punto sobre la recta es un múltiplo de a. El problema es calcular el coeficiente $\hat{x}$. Todo lo que se requiere es el hecho geométrico de que ***la recta desde b al punto más próximo $p = \hat{x}a$ es perpendicular al vector a*:**

$$(b-\hat{x}a) \perp a, \quad \text{o bien} \quad a^{\mathrm{T}}(b-\hat{x}a) = 0, \quad \text{o bien} \quad \hat{x} = \frac{a^{\mathrm{T}}b}{a^{\mathrm{T}}a}. \quad (4)$$

Así se obtiene la fórmula para el número $\hat{x}$ y la proyección p:

3H La proyección del vector b sobre la recta en la dirección de a es $p = \hat{x}a$:

Proyección sobre una recta $$p = \hat{x}a = \frac{a^{\mathrm{T}}b}{a^{\mathrm{T}}a}a. \quad (5)$$

Esto permite volver a trazar la figura 3.5 con una fórmula correcta para p (véase la figura 3.7).

Lo anterior conduce a la desigualdad de Schwarz en la ecuación (6), que es la desigualdad más importante en matemáticas. Un caso especial es el hecho de que las medias aritméticas $\frac{1}{2}(x+y)$ son más grandes que las medias geométricas $\sqrt{xy}$. (Esto es equivalente –consulte el problema 1 al final de esta sección– a la desigualdad del triángulo para vectores.) Parece que la desigualdad de Schwarz surge casi accidentalmente de la afirmación

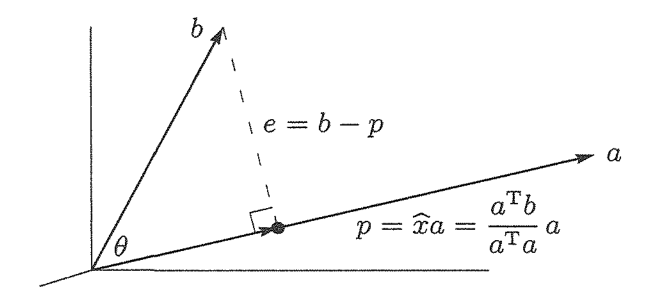


Figura 3.7 Proyección p de b sobre a, con $\cos\theta = \dfrac{Op}{Ob} = \dfrac{a^{\mathrm{T}}b}{\|a\|\,\|b\|}$.

de que $\|e\|^2 = \|b-p\|^2$ en la figura 3.7 no puede ser negativa:

$$\left\|b - \frac{a^{\mathrm{T}}b}{a^{\mathrm{T}}a}a\right\|^2 = b^{\mathrm{T}}b - 2\frac{(a^{\mathrm{T}}b)^2}{a^{\mathrm{T}}a} + \left(\frac{a^{\mathrm{T}}b}{a^{\mathrm{T}}a}\right)^2 a^{\mathrm{T}}a = \frac{(b^{\mathrm{T}}b)(a^{\mathrm{T}}a) - (a^{\mathrm{T}}b)^2}{(a^{\mathrm{T}}a)} \geq 0.$$

Lo anterior indica que $(b^{\mathrm{T}}b)(a^{\mathrm{T}}a) \geq (a^{\mathrm{T}}b)^2$, y luego se toman las raíces cuadradas:

3I Todos los vectores satisfacen la ***desigualdad de Schwarz***, que es $|\cos\theta| \leq 1$ en $\mathbf{R}^n$:

Desigualdad de Schwarz $$|a^{\mathrm{T}}b| \leq \|a\|\,\|b\|. \quad (6)$$

Según la fórmula (2), la razón entre $a^{\mathrm{T}}b$ y $\|a\|\,\|b\|$ es exactamente $|\cos\theta|$. Debido a que todos los cosenos están en el intervalo $-1 \leq \cos\theta \leq 1$, esto constituye otra demostración de la ecuación (6): *la desigualdad de Schwarz es lo mismo que* $|\cos\theta| \leq 1$. De alguna manera es una demostración que se entiende más fácilmente debido a que los cosenos son bastante conocidos. Cualquier demostración está bien en $\mathbf{R}^n$, aunque debe observar que la aquí proporcionada provino directamente del cálculo de $\|b-p\|^2$. Esto permanece no negativo cuando se introducen nuevas posibilidades para las longitudes y los productos internos. El nombre de Cauchy también está ligado a esta desigualdad $|a^{\mathrm{T}}b| \leq \|a\|\,\|b\|$, ¡y los rusos se refieren a ella como la desigualdad de Cauchy-Schwarz-Buniakowsky! Los historiadores expertos en matemáticas parecen coincidir en que la reclamación de Buniakowsky es legítima.

Una observación final sobre $|a^{\mathrm{T}}b| \leq \|a\|\,\|b\|$. *La igualdad se cumple si y sólo si b es un múltiplo de a*. El ángulo es $\theta = 0°$ o $\theta = 180°$ y el coseno es 1 o –1. En este caso b es idéntico a su proyección p, y la distancia entre b y la recta es cero.

Ejemplo 1 Proyectar $b = (1, 2, 3)$ sobre la recta que pasa por $a = (1, 1, 1)$ para obtener $\hat{x}$ y p:

$$\hat{x} = \frac{a^{\mathrm{T}}b}{a^{\mathrm{T}}a} = \frac{6}{3} = 2.$$

La proyección es $p = \hat{x}a = (2, 2, 2)$. El ángulo entre a y b tiene

$$\cos\theta = \frac{\|p\|}{\|b\|} = \frac{\sqrt{12}}{\sqrt{14}} \quad \text{y también} \quad \cos\theta = \frac{a^{\mathrm{T}}b}{\|a\|\,\|b\|} = \frac{6}{\sqrt{3}\sqrt{14}}.$$

La desigualdad de Schwarz $|a^{\mathrm{T}}b| \leq \|a\|\,\|b\|$ es $6 \leq \sqrt{3}\sqrt{14}$. Si 6 se escribe como $\sqrt{36}$, esto es lo mismo que $\sqrt{36} \leq \sqrt{42}$. El coseno es menor que 1, ya que b no es paralelo a a.

Matriz proyección de rango 1

La proyección de b sobre la recta que pasa por a está en $p = a(a^Tb/a^Ta)$. Esta es nuestra fórmula $p = \hat{x}a$, aunque aquí se ha escrito con un leve giro: El vector a aparece antes que el número $\hat{x} = a^Tb/a^Ta$. Detrás de este cambio aparentemente trivial hay una razón. La proyección sobre una recta es llevada a cabo por una ***matriz proyección P***, y escrita en este nuevo orden puede verse de qué se trata. *P es la matriz que multiplica a b y produce p*:

$$P = a\frac{a^Tb}{a^Ta} \quad \text{de modo que la matriz proyección es} \quad P = \frac{aa^T}{a^Ta}. \tag{7}$$

Esto es una columna multiplicada por un renglón —una matriz cuadrada— dividida entre el número a^Ta.

Ejemplo 2 La matriz que proyecta sobre la recta que pasa por $a = (1, 1, 1)$ es

$$P = \frac{aa^T}{a^Ta} = \frac{1}{3}\begin{bmatrix}1\\1\\1\end{bmatrix}\begin{bmatrix}1 & 1 & 1\end{bmatrix} = \begin{bmatrix}\frac{1}{3} & \frac{1}{3} & \frac{1}{3}\\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}\\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}\end{bmatrix}.$$

Esta matriz tiene dos propiedades que serán consideradas como típicas de las proyecciones:

1. **P es una matriz simétrica.**
2. **Su cuadrado es ella misma: $P^2 = P$.**

P^2b es la proyección de Pb, ¡y Pb ya está en la recta! Así, $P^2b = Pb$. Esta matriz P también constituye un gran ejemplo de los cuatro subespacios fundamentales:

El espacio columna consta de la recta que pasa por $a = (1, 1, 1)$.
El espacio nulo consta del plano perpendicular a a.
El rango es $r = 1$.

Toda columna es un múltiplo de a, así como también lo es $Pb = \hat{x}a$. Los vectores que proyectan hacia $p = 0$ son especialmente importantes. Satisfacen $a^Tb = 0$ –son perpendiculares a a y su componente a lo largo de la recta es cero. Están en el espacio nulo = plano perpendicular.

En realidad, este ejemplo es demasiado perfecto. Tiene el espacio nulo ortogonal al espacio columna, que se ha vuelto loco. El espacio nulo debe ser ortogonal al *espacio renglón*. Pero como P es simétrica, sus espacios renglón y columna son el mismo.

Observación sobre la escalación La matriz proyección aa^T/a^Ta es igual a a si se duplica:

$$a = \begin{bmatrix}2\\2\\2\end{bmatrix} \quad \text{proporciona} \quad P = \frac{1}{12}\begin{bmatrix}2\\2\\2\end{bmatrix}\begin{bmatrix}2 & 2 & 2\end{bmatrix} = \begin{bmatrix}\frac{1}{3} & \frac{1}{3} & \frac{1}{3}\\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}\\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}\end{bmatrix} \quad \text{como antes.}$$

La recta que pasa por a es la misma, y eso es todo lo que le importa a la matriz proyección. Si la longitud de a es 1, el denominador es $a^Ta = 1$ y la matriz es justo $P = aa^T$.

Ejemplo 3 Proyectar sobre la "dirección θ" en el plano x-y. La recta pasa por $a = (\cos\theta, \operatorname{sen}\theta)$ y la matriz es simétrica con $P^2 = P$:

$$P = \frac{aa^T}{a^Ta} = \frac{\begin{bmatrix}c\\s\end{bmatrix}\begin{bmatrix}c & s\end{bmatrix}}{\begin{bmatrix}c & s\end{bmatrix}\begin{bmatrix}c\\s\end{bmatrix}} = \begin{bmatrix}c^2 & cs\\ cs & s^2\end{bmatrix}.$$

Aquí, c es $\cos\theta$, s es $\operatorname{sen}\theta$ y $c^2 + s^2 = 1$ en el denominador. Esta matriz P se descubrió en la sección 2.6 sobre transformaciones lineales. Ahora, P se conoce en cualquier número de dimensiones. Se recalca que produce la proyección p:

Para proyectar b sobre a, es necesario multiplicar la matriz proyección P: p = Pb.

Traspuestas de productos internos

Por último, los productos internos se relacionan con A^T. Hasta el momento, A^T es simplemente la reflexión de A a través de la diagonal principal: los renglones de A se convierten en las columnas de A^T, y viceversa. El elemento en el renglón i, columna j de A es el elemento (j, i) de A:

Traspuesta por reflexión $\quad (A^T)_{ij} = (A)_{ji}$.

Hay un significado más profundo de A^T. Su estrecha relación con los productos internos constituye una nueva y mucho más "abstracta" definición de la traspuesta:

3J La traspuesta A^T puede definirse mediante la siguiente propiedad: el producto interno de Ax con y es igual al producto interno de x con A^Ty. Formalmente, esto significa simplemente que

$$(Ax)^Ty = x^TA^Ty = x^T(A^Ty). \tag{8}$$

Esta definición constituye otra (mejor) manera de comprobar la fórmula $(AB)^T = B^TA^T$. La ecuación (8) se aplica dos veces:

Primero se mueve A y luego se mueve B $\quad (ABx)^Ty = (Bx)^T(A^Ty) = x^T(B^TA^Ty)$.

La traspuesta se da vuelta en orden inverso sobre el lado derecho, justo como hacen las inversas en la fórmula $(AB)^{-1} = B^{-1}A^{-1}$. Se hace mención nuevamente que estas dos fórmulas se encuentran para producir la extraordinaria combinación $(A^{-1})^T = (A^T)^{-1}$.

Conjunto de problemas 3.2

1. *a*) Dados dos números positivos cualesquiera x y y, escoja el vector b igual a $(\sqrt{x}, \sqrt{y})$, y escoja $a = (\sqrt{y}, \sqrt{x})$. Aplique la desigualdad de Schwarz para comparar la media aritmética $\frac{1}{2}(x + y)$ con la media geométrica $\sqrt{xy}$.
   b) Suponga que se empieza con un vector que va del origen al punto x, y que luego se suma un vector de longitud $\|y\|$ que une x con $x + y$. El tercer lado del triángulo va del origen a $x + y$. *La desigualdad del triángulo establece que esta distancia no puede ser mayor que la suma de las dos primeras distancias*:

$$\|x + y\| \le \|x\| + \|y\|.$$

Después de elevar al cuadrado ambos miembros, y desarrollar $(x+y)^T(x+y)$, reduzca lo anterior a la desigualdad de Schwarz.

2. Compruebe que la longitud de la proyección en la figura 3.7 es $\|p\| = \|b\| \cos\theta$, aplicando la fórmula (5).

3. ¿Qué múltiplo de $a = (1, 1, 1)$ está más próximo al punto $b = (2, 4, 4)$? También encuentre el punto más próximo a a sobre la recta que pasa por b.

4. Explique por qué la desigualdad de Schwarz se vuelve igualdad cuando a y b están sobre la misma recta que pasa por el origen, y sólo en ese caso. ¿Qué ocurre si están en lados opuestos del origen?

5. En n dimensiones, ¿qué ángulo forma el vector $(1, 1, \ldots, 1)$ con los ejes de coordenadas? ¿Cuál es la matriz proyección P sobre ese vector?

6. La desigualdad de Schwarz puede demostrarse en una línea si a y b se normalizan delante del tiempo para que sean vectores unitarios:

$$|a^Tb| = \left|\sum a_jb_j\right| \le \sum |a_j||b_j| \le \sum \frac{|a_j|^2 + |b_j|^2}{2} = \frac{1}{2} + \frac{1}{2} = \|a\|\,\|b\|.$$

¿Cuál de los problemas previos justifica el paso de en medio?

7. Escoja el vector correcto b en la desigualdad de Schwarz para demostrar que

$$(a_1 + \cdots + a_n)^2 \le n\left(a_1^2 + \cdots + a_n^2\right).$$

¿Cuándo se cumple la igualdad?

8. La molécula de metano CH_4 está dispuesta como si el átomo de carbono estuviese en el centro de un tetraedro regular con cuatro átomos de hidrógeno en los vértices. Si los vértices están en (0, 0, 0), (1, 1, 0), (1, 0, 1), y (0, 1, 1) —observe que la longitud de todas las seis aristas es $\sqrt{2}$, de modo que el tetraedro es regular—, ¿cuál es el coseno del ángulo formado por los rayos que van del centro $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$ a los vértices? (El ángulo de enlace en sí mide aproximadamente 109.5°, un viejo conocido de los químicos.)

9. Eleve al cuadrado la matriz $P = aa^T/a^Ta$, que proyecta sobre una recta, y demuestre que $P^2 = P$. (¡Observe el número a^Ta en medio de la matriz aa^Taa^T!)

10. La matriz proyección P, ¿es invertible? ¿Por qué?

11. *a*) Encuentre la matriz proyección P_1 sobre la recta que pasa por $a = \begin{bmatrix}1\\3\end{bmatrix}$ y también la matriz P_2 que proyecta sobre la recta perpendicular a a.

    b) Calcule $P_1 + P_2$, P_1P_2, y explique.

12. Encuentre la matriz que proyecta todo punto en el plano sobre la recta $x + 2y = 0$.

13. Demuestre que *la traza* de $P = aa^T/a^Ta$ —que es la suma de sus elementos diagonales— siempre es igual a 1.

14. ¿Cuál matriz P proyecta todo punto de $\mathbf{R}^3$ sobre la recta de intersección de los planos $x + y + t = 0$ y $x - t = 0$?

15. Demuestre que la longitud de Ax es igual a la longitud de A^Tx si $AA^T = A^TA$.

16. Suponga que P es la matriz proyección sobre la recta que pasa por a.

    a) ¿Por qué el producto interno de x con Py es igual al producto interno de Px con y?

    b) ¿Son iguales los dos ángulos? Encuentre sus cosenos si $a = (1, 1, -1)$, $x = (2, 0, 1)$, y $y = (2, 1, 2)$.

    c) ¿Por qué el producto interno de Px con Py siempre es el mismo? ¿Cuál es el ángulo entre Px y Py?

En los problemas 17 a 26 se piden proyecciones sobre rectas. También los errores $e = b - p$, y las matrices P.

17. Proyecte el vector b sobre la recta que pasa por a. Compruebe que e es perpendicular a a:

$$a)\quad b = \begin{bmatrix}1\\2\\2\end{bmatrix} \quad \text{y} \quad a = \begin{bmatrix}1\\1\\1\end{bmatrix}. \qquad b)\quad b = \begin{bmatrix}1\\3\\1\end{bmatrix} \quad \text{y} \quad a = \begin{bmatrix}-1\\-3\\-1\end{bmatrix}.$$

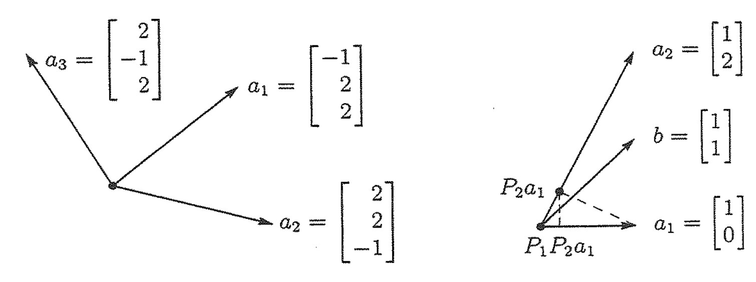
18. *Trace* la proyección de b sobre a y también calcúlela a partir de $p = \hat{x}a$:

$$a)\quad b = \begin{bmatrix}\cos\theta\\ \operatorname{sen}\theta\end{bmatrix} \quad \text{y} \quad a = \begin{bmatrix}1\\0\end{bmatrix}. \qquad b)\quad b = \begin{bmatrix}1\\1\end{bmatrix} \quad \text{y} \quad a = \begin{bmatrix}1\\-1\end{bmatrix}.$$

19. En el problema 17, encuentre la matriz proyección $P = aa^T/a^Ta$ sobre la recta que pasa por cada vector a. En ambos casos, compruebe que $P^2 = P$. Multiplique Pb en cada caso para calcular la proyección p.

20. Construya las matrices proyección P_1 y P_2 sobre las rectas que pasan por cada vector a del problema 18. ¿Es cierto que $(P_1 + P_2)^2 = P_1 + P_2$? Esto *sería* cierto si $P_1P_2 = 0$.

Para los problemas 21 a 26, consulte las figuras que se muestran a continuación.



Problemas 21 a 23 | Problemas 24 a 26

21. Construya las matrices proyección aa^T/a^Ta sobre las rectas que pasan por $a_1 = (-1, 2, 2)$ y $a_2 = (2, 2, -1)$. Multiplique estas matrices proyección, y explique por qué su producto P_1P_2 es lo que es.

22. Proyecte $b = (1, 0, 0)$ sobre las rectas que pasan por a_1 y a_2 del problema 21 y también sobre $a_3 = (2, -1, 2)$. Sume las tres proyecciones $p_1 + p_2 + p_3$.

23. Continuando con los problemas 21 y 22, encuentre la matriz proyección P_3 sobre $a_3 = (2, -1, 2)$. Compruebe que $P_1 + P_2 + P_3 = I$. ¡La base a_1, a_2, a_3 es ortogonal!

24. Proyecte el vector $b = (1, 1)$ sobre las rectas que pasan por $a_1 = (1, 0)$ y $a_2 = (1, 2)$. Trace las proyecciones p_1 y p_2 y sume $p_1 + p_2$. La suma de las proyecciones no es b porque las as no son ortogonales.

25. En el problema 24, la proyección de b sobre el *plano* de a_1 y a_2 es igual a b. Encuentre $P = A(A^TA)^{-1}A^T$ para $A = \begin{bmatrix} a_1 & a_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$.

26. Proyecte $a_1 = (1, 0)$ sobre $a_2 = (1, 2)$. Luego proyecte el resultado de vuelta sobre a_1. Trace estas proyecciones y multiplique las matrices proyección P_1P_2: ¿Ésta es una proyección?

3.3 PROYECCIONES Y MÍNIMOS CUADRADOS

Hasta el momento, $Ax = b$ tiene una solución o no. Si b no está en el espacio columna $C(A)$, entonces el sistema es inconsistente y la eliminación gaussiana fracasa. Casi seguramente, esta falla ocurre cuando hay muchas ecuaciones y sólo una incógnita:

Más ecuaciones que incógnitas: ¿no hay solución?
$$2x = b_1 \qquad 3x = b_2 \qquad 4x = b_3.$$

Este sistema es resoluble cuando la razón entre b_1, b_2 y b_3 es 2:3:4. La solución x existe sólo si b está sobre la misma recta que la columna $a = (2, 3, 4)$.

A pesar de su irresolubilidad, las ecuaciones inconsistentes se presentan todo el tiempo en la práctica. ¡Es necesario resolverlas! Una posibilidad es determinar x con base en una parte de ese sistema e ignorar el resto; esto es difícil de justificar si todas las m ecuaciones vienen de la misma fuente. En vez de esperar que en algunas ecuaciones no haya error y que en otras el error sea grande, es mucho mejor *escoger la x que minimiza un error promedio E en las m ecuaciones.*

El "promedio" más conveniente proviene de la *suma de los cuadrados*:

Error al cuadrado $$E^2 = (2x - b_1)^2 + (3x - b_2)^2 + (4x - b_3)^2.$$

Si hay una solución exacta, el error mínimo es $E = 0$. En el caso más probable en que b no sea proporcional a a, la gráfica de E^2 es una parábola. El error mínimo está en el punto más bajo, donde la derivada es cero:

$$\frac{dE^2}{dx} = 2[(2x - b_1)2 + (3x - b_2)3 + (4x - b_3)4] = 0.$$

Al despejar x, la solución por mínimos cuadrados de este sistema modelo $ax = b$ se denota por $\widehat{x}$:

Solución por mínimos cuadrados $$\widehat{x} = \frac{2b_1 + 3b_2 + 4b_3}{2^2 + 3^2 + 4^2} = \frac{a^Tb}{a^Ta}.$$

En el numerador se reconoce a^Tb, y en el denominador, a^Ta.

El caso general es el mismo. $ax = b$ se "resuelve" al minimizar

$$E^2 = \|ax - b\|^2 = (a_1x - b_1)^2 + \cdots + (a_mx - b_m)^2.$$

La derivada de E^2 es cero en el punto $\widehat{x}$, si

$$(a_1\widehat{x} - b_1)a_1 + \cdots + (a_m\widehat{x} - b_m)a_m = 0.$$

Se está minimizando la distancia de b a la recta que pasa por a, y en cálculo se obtiene la misma respuesta, $\widehat{x} = (a_1b_1 + \cdots + a_mb_m)/(a_1^2 + \cdots + a_m^2)$, que se obtuvo antes con la geometría:

3K La solución por mínimos cuadrados de un problema $ax = b$ en una incógnita es
$$\widehat{x} = \frac{a^Tb}{a^Ta}.$$

Puede darse cuenta de que seguimos regresando a la interpretación geométrica de un problema de mínimos cuadrados: minimizar una distancia. Al igualar a cero la derivada de E^2, el cálculo confirma la geometría de la sección previa. *El vector error e que une b con p debe ser perpendicular a a:*

Ortogonalidad de a y e $$a^T(b - \widehat{x}a) = a^Tb - \frac{a^Tb}{a^Ta}a^Ta = 0.$$

Como observación al margen, observe el caso degenerado $a = 0$. Todos los múltiplos de a son cero, y la recta es sólo un punto. En consecuencia, $p = 0$ es el único candidato para la proyección. Sin embargo, la fórmula para $\widehat{x}$ se convierte en una expresión 0/0 sin sentido, y correctamente refleja el hecho de que $\widehat{x}$ está completamente indeterminado. Todos los valores de x proporcionan el mismo error $E = \|0x - b\|$, de modo que E^2 es una recta horizontal en vez de una parábola. La "seudoinversa" asigna el valor definitivo $\widehat{x} = 0$, que es una elección más "simétrica" que cualquier otra.

Problemas de mínimos cuadrados con varias variables

Ahora ya estamos listos para dar un paso importante: ***proyectar b sobre un subespacio*** –en vez de hacerlo simplemente sobre una recta. Este problema se presenta de $Ax = b$ cuando A es una matriz de m por n. En vez de una columna y una incógnita x, ahora la matriz tiene n columnas. El número m de observaciones sigue siendo mayor que el número n de incógnitas, por lo que es de esperar que $Ax = b$ sea inconsistente. *Quizá no haya una elección de x que se ajuste perfectamente a los datos b.* En otras palabras, probablemente el vector b no sea una combinación de las columnas de A; está fuera del espacio columna.

De nuevo el problema es escoger $\widehat{x}$ con la finalidad de minimizar el error, y de nuevo esta minimización se hace en el sentido de mínimos cuadrados. El error es $E = \|Ax - b\|$, ***y es exactamente la distancia de b al punto Ax en el espacio columna.*** Buscar la solución por mínimos cuadrados $\widehat{x}$, que minimiza E, es lo mismo que ubicar el punto $p = A\widehat{x}$ que está más próximo a b que cualquier otro punto en el espacio columna.

Para determinar $\widehat{x}$ puede usarse geometría o cálculo. En n dimensiones es preferible utilizar geometría; p debe ser la "proyección de b sobre el espacio columna". ***El vector error $e = b - A\widehat{x}$ debe ser perpendicular a ese espacio*** (véase la figura 3.8). Encontrar $\widehat{x}$ y la proyección $p = A\widehat{x}$ es tan fundamental que se hace de dos maneras:

1. Todos los vectores perpendiculares al espacio columna están en el *espacio nulo izquierdo*. Así, el vector error $e = b - A\widehat{x}$ debe estar en el espacio nulo de A^T:

$$A^T(b - A\widehat{x}) = 0 \quad \text{o bien} \quad A^TA\widehat{x} = A^Tb.$$

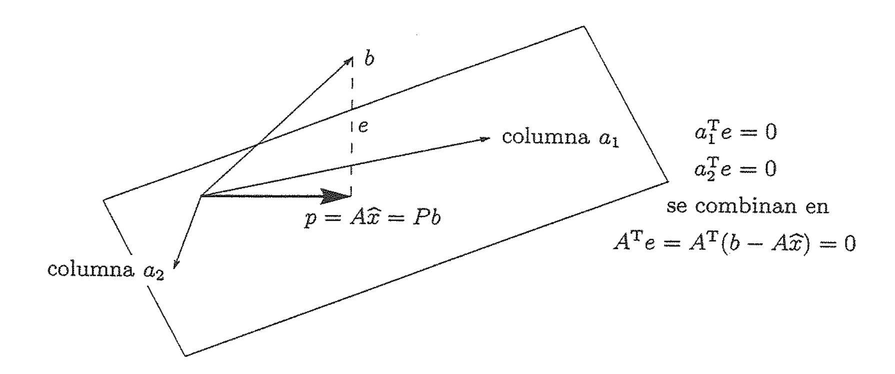


Figura 3.8 Proyección sobre el espacio columna de una matriz de 3 por 2.

2. El vector error debe ser perpendicular a *cada columna* $a_1, \ldots, a_n$ de A:

$$\begin{matrix} a_1^{\mathrm{T}}(b - A\widehat{x}) = 0 \\ \vdots \\ a_n^{\mathrm{T}}(b - A\widehat{x}) = 0 \end{matrix} \quad \text{o bien} \quad \begin{bmatrix} a_1^{\mathrm{T}} \\ \vdots \\ a_n^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} b - A\widehat{x} \end{bmatrix} = 0.$$

Esto es nuevamente $A^{\mathrm{T}}(b - A\widehat{x}) = 0$ y $A^{\mathrm{T}}A\widehat{x} = A^{\mathrm{T}}b$. El método del cálculo consiste en tomar derivadas parciales de $E^2 = (Ax - b)^{\mathrm{T}}(Ax - b)$. Así se obtiene lo mismo que $2A^{\mathrm{T}}Ax - 2A^{\mathrm{T}}b = 0$. La forma más rápida es ***multiplicar la ecuación irresoluble $Ax = b$ por A^{T}***. Con todos estos métodos equivalentes se obtiene una matriz de coeficientes cuadrada $A^{\mathrm{T}}A$. Es simétrica (¡su traspuesta no es AA^{T}!) y es la matriz fundamental de este capítulo.

En estadística, las ecuaciones $A^{\mathrm{T}}A\widehat{x} = A^{\mathrm{T}}b$ se conocen como ***ecuaciones normales***.

3L Cuando $Ax = b$ es inconsistente, su solución por mínimos cuadrados minimiza $\|Ax - b\|^2$:

$$\textbf{Ecuaciones normales} \quad A^{\mathrm{T}}A\widehat{x} = A^{\mathrm{T}}b. \tag{1}$$

¡$A^{\mathrm{T}}A$ es invertible exactamente cuando las columnas de A son linealmente independientes! Así,

$$\textbf{Mejor estimación } \widehat{x} \quad \widehat{x} = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}b. \tag{2}$$

La proyección de b sobre el espacio columna es el punto más próximo $A\widehat{x}$:

$$\textbf{Proyección} \quad p = A\widehat{x} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}b. \tag{3}$$

Se escoge un ejemplo en donde nuestra intuición es tan buena como las fórmulas:

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \quad \begin{matrix} Ax = b & \text{no tiene solución} \\ A^{\mathrm{T}}A\widehat{x} = A^{\mathrm{T}}b & \text{proporciona la mejor } x. \end{matrix}$$

Ambas columnas terminan con un cero, de modo que $C(A)$ es el plano x-y dentro del espacio tridimensional. La proyección de $b = (4, 5, 6)$ es $p = (4, 5, 0)$; las componentes x y y

permanecen igual pero $z = 6$ desaparece. Este hecho lo confirman las ecuaciones normales:

$$A^{\mathrm{T}}A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 5 & 13 \end{bmatrix}.$$

$$\widehat{x} = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}b = \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

$$\textbf{Proyección} \quad p = A\widehat{x} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix}.$$

En este caso especial, lo mejor que puede hacerse es resolver las dos primeras ecuaciones de $Ax = b$. Así, $\widehat{x}_1 = 2$ y $\widehat{x}_2 = 1$. Es seguro que el error en la ecuación $0x_1 + 0x_2 = 6$ es el 6.

Observación 1 Suponga que b está realmente *en* el espacio columna de A: es una combinación $b = Ax$ de las columnas. Entonces, la proyección de b sigue siendo b:

$$\textbf{b en el espacio columna} \quad p = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}Ax = Ax = b.$$

El punto p más próximo es justo b mismo, lo cual es evidente.

Observación 2 En el otro extremo, suponga que b es *perpendicular* a cada columna, de modo que $A^{\mathrm{T}}b = 0$. En este caso b se proyecta al vector cero:

$$\textbf{b en el espacio nulo izquierdo} \quad p = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}b = A(A^{\mathrm{T}}A)^{-1}0 = 0.$$

Observación 3 Cuando A es cuadrada e invertible, el espacio columna es todo el espacio. Cada vector se proyecta en sí mismo, p es igual a b, y $\widehat{x} = x$:.

$$\textbf{Si A es invertible} \quad p = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}b = AA^{-1}(A^{\mathrm{T}})^{-1}A^{\mathrm{T}}b = b.$$

Este es el único caso en que es posible separar $(A^{\mathrm{T}}A)^{-1}$, *y escribirla como* $A^{-1}(A^{\mathrm{T}})^{-1}$. Cuando A es rectangular esto no es posible.

Observación 4 Suponga que A sólo tiene una columna, que contiene a a. Entonces la matriz $A^{\mathrm{T}}A$ es el número $a^{\mathrm{T}}a$ y $\widehat{x}$ es $a^{\mathrm{T}}b/a^{\mathrm{T}}a$. Se regresa a la primera fórmula.

La matriz producto cruz $A^{\mathrm{T}}A$

La matriz $A^{\mathrm{T}}A$ es ciertamente simétrica. Su traspuesta es $(A^{\mathrm{T}}A)^{\mathrm{T}} = A^{\mathrm{T}}A^{\mathrm{TT}}$, que es $A^{\mathrm{T}}A$ de nuevo. Su elemento i, j (y el elemento j, i) es el producto interno de la columna i de A con la columna j de A. La cuestión clave es la invertibilidad de $A^{\mathrm{T}}A$, y por fortuna

$A^{\mathrm{T}}A$ tiene el mismo espacio nulo que A.

Ciertamente, si $Ax = 0$, entonces $A^{\mathrm{T}}Ax = 0$. Los vectores x en el espacio nulo de A también están en el espacio nulo de $A^{\mathrm{T}}A$. Para ir en la otra dirección, debe empezarse suponiendo

que $A^TAx = 0$, y tomar el producto interno con x para demostrar que $Ax = 0$:

$$x^TA^TAx = 0, \quad \text{o bien,} \quad \|Ax\|^2 = 0, \quad \text{o bien,} \quad Ax = 0.$$

Los dos espacios nulos son idénticos. En particular, si las columnas de A son independientes (y sólo $x = 0$ está en su espacio nulo), entonces lo mismo es cierto para A^TA:

3M Si A tiene columnas independientes, entonces A^TA *es cuadrada, simétrica e invertible.*

Después se demostrará que A^TA también es positiva definida (todos los pivotes y los valores característicos son positivos).

Este caso es por mucho el más común e importante. La independencia no es tan difícil en un espacio m-dimensional si $m > n$. Esto se supone de aquí en adelante.

Matrices proyección

Se ha demostrado que el punto más próximo a b es $p = A(A^TA)^{-1}A^Tb$. *Esta fórmula expresa en términos matriciales la construcción de una recta perpendicular de b al espacio columna de A.* La matriz que proporciona P es una matriz proyección, denotada por P:

$$\textbf{Matriz proyección} \qquad P = A(A^TA)^{-1}A^T. \tag{4}$$

Esta matriz proyecta cualquier vector b sobre el espacio columna de A.* En otras palabras, $p = Pb$ es la componente de b en el espacio columna, y el error $e = b - Pb$ es la componente en el complemento ortogonal. (¡$I - P$ también es una matriz proyección! Proyecta b sobre el complemento ortogonal, y la proyección es $b - Pb$.)

En breve, se tiene una fórmula matricial para separar cualquier b en dos componentes perpendiculares. Pb está en el espacio columna $C(A)$, y la otra componente $(I - P)b$ está en el espacio nulo izquierdo $N(A^T)$, que es ortogonal al espacio columna.

Las matrices proyección pueden entenderse geométrica y algebraicamente.

3N La matriz proyección $P = A(A^TA)^{-1}A^T$ tiene dos propiedades fundamentales:

i) Es igual a su cuadrado: $P^2 = P$.

ii) Es igual a su traspuesta: $P^T = P$.

Al revés, cualquier matriz simétrica con $P^2 = P$ representa una proyección.

Demostración Es fácil ver por qué $P^2 = P$. Si se empieza con cualquier b, entonces Pb está en el subespacio sobre el que se está proyectando. ***Cuando se vuelve a proyectar nada cambia.*** El vector Pb ya está en el subespacio, y $P(Pb)$ sigue siendo Pb. En otras palabras, $P^2 = P$. Dos o tres o cincuenta proyecciones proporcionan el mismo punto p que la primera proyección:

$$P^2 = A(A^TA)^{-1}A^TA(A^TA)^{-1}A^T = A(A^TA)^{-1}A^T = P.$$

*Puede haber un riesgo de confusión con las matrices permutación, que también se denotan por P, aunque el riesgo debe ser pequeño, y trataremos que nunca aparezcan ambas en la misma página.

Para demostrar que P también es simétrica, se toma su traspuesta. Las traspuestas se multiplican en orden invertido, y se usa la simetría de $(A^TA)^{-1}$, para regresar a P:

$$P^T = (A^T)^T((A^TA)^{-1})^TA^T = A((A^TA)^T)^{-1}A^T = A(A^TA)^{-1}A^T = P.$$

Para la conversa, a partir de $P^2 = P$ y $P^T = P$ es necesario deducir que Pb **es la proyección ortogonal de b sobre el espacio columna de P.** El vector error $b - Pb$ es *ortogonal al espacio.* Para cualquier vector Pc en el espacio, el producto interior es cero:

$$(b - Pb)^TPc = b^T(I - P)^TPc = b^T(P - P^2)c = 0.$$

Así, $b - Pb$ es ortogonal al espacio, y Pb es la proyección sobre el espacio columna. ■

Ejemplo 1 Suponga que A es verdaderamente invertible. Si es de 4 por 4, entonces sus cuatro columnas son independientes y su espacio columna es todo $\mathbf{R}^4$. ¿Cuál es la proyección *sobre todo el espacio*? Es la matriz identidad.

$$P = A(A^TA)^{-1}A^T = AA^{-1}(A^T)^{-1}A^T = I. \tag{5}$$

La matriz identidad es simétrica, $I^2 = I$, y el error $b - Ib$ es cero.

La cuestión con todos los otros ejemplos es que lo que ocurrió en la ecuación (5) *no se permite*. Para repetir: No es posible invertir las partes por separado A^T y A cuando estas matrices son rectangulares. La que es invertible es la matriz A^TA.

Ajuste de datos por mínimos cuadrados

Suponga que se realiza una serie de experimentos, y se espera que la salida b sea una función lineal de la entrada t. Se busca una ***línea recta*** $b = C + Dt$. Por ejemplo:

1. En instantes diferentes se mide la distancia a un satélite en su camino a Marte. En este caso t es el tiempo y b es la distancia. A menos que el motor esté encendido o la gravedad sea intensa, el satélite se moverá casi a velocidad constante v: $b = b_0 + vt$.
2. En una estructura se hace variar la carga, y se mide el movimiento que esto produce. En este experimento t es la carga y b es la lectura en el medidor de deformación. A menos que la carga sea tan grande que el material se vuelva plástico, en la teoría de la elasticidad es normal una relación lineal $b = C + Dt$.
3. El costo de producción de t libros como éste es casi lineal, $b = C + Dt$, con la edición y la composición en C y luego la impresión y la encuadernación en D. Así, C es el costo de preparación y D es el costo por cada libro adicional.

¿Cómo calcular C y D? Si no hay error experimental, entonces dos mediciones de b determinan la recta $b = C + Dt$. Pero si hay error, es necesario estar preparado para "promediar" los experimentos y encontrar una recta óptima. ¡Esta recta no debe confundirse con la recta que pasa por a sobre la cual se proyectó b en la sección previa! De hecho, ya que hay dos incógnitas por determinar: C y D, ahora se proyecta sobre un subespacio *bidimensional*. Un experimento perfecto debe proporcionar una C y una D perfectas:

$$\begin{aligned} C + Dt_1 &= b_1 \\ C + Dt_2 &= b_2 \\ &\vdots \\ C + Dt_m &= b_m. \end{aligned} \tag{6}$$

Este es un sistema *sobredeterminado*, con m ecuaciones y sólo dos incógnitas. Si hay errores presentes, el sistema no tiene solución. A tiene dos columnas, y $x = (C, D)$:

$$\begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \quad \text{o bien,} \quad Ax = b. \tag{7}$$

La mejor solución $(\widehat{C}, \widehat{D})$ es la $\widehat{x}$ que minimiza el error al cuadrado E^2:

$$\textbf{Minimizar} \quad E^2 = \|b - Ax\|^2 = (b_1 - C - Dt_1)^2 + \cdots + (b_m - C - Dt_m)^2.$$

El vector $p = A\widehat{x}$ está lo más próximo posible a b. De todas las rectas $b = C + Dt$, se está eligiendo la que mejor se ajusta a los datos (véase la figura 3.9). En la gráfica, los errores son las ***distancias verticales*** $b - C - Dt$ a la recta (¡no son distancias perpendiculares!) Son las distancias verticales las que se elevan al cuadrado, se suman, y minimizan.

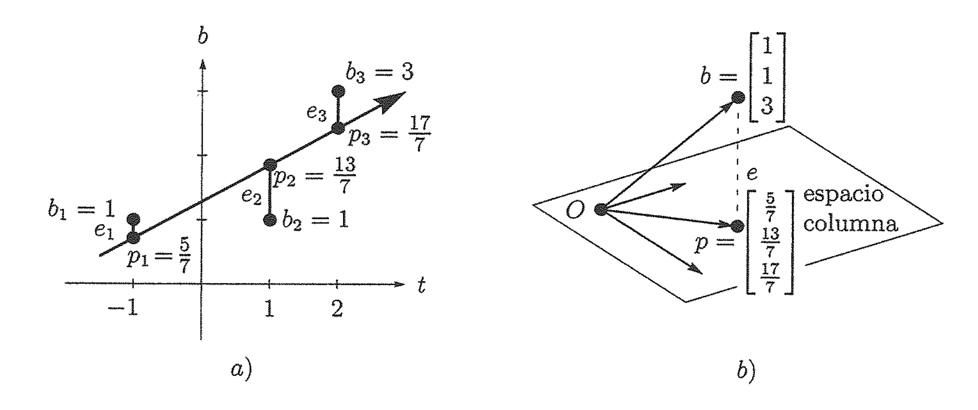


Figura 3.9 La aproximación por una recta corresponde a la proyección p de b.

Ejemplo 2 En la figura 3.9*a* se muestran tres mediciones b_1, b_2, b_3:

$$b = 1 \text{ en } t = -1, \qquad b = 1 \text{ en } t = 1, \qquad b = 3 \text{ en } t = 2.$$

Observe que no se requiere que los valores $t = -1, 1, 2$ sean equidistantes. El primer paso es *escribir las ecuaciones que se **cumplirían** si una recta pudiera pasar por todos los tres puntos.*

Entonces $C + Dt$ coincidiría exactamente con b:

$$Ax = b \quad \text{es} \quad \begin{aligned} C - D &= 1 \\ C + D &= 1 \\ C + 2D &= 3 \end{aligned} \quad \text{o bien} \quad \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}.$$

Si estas ecuaciones $Ax = b$ pudieran resolverse, no habría errores. No es posible resolverlas porque los puntos no están en una recta. En consecuencia, se resuelven por mínimos cuadrados:

$$A^{\mathrm{T}}A\widehat{x} = A^{\mathrm{T}}b \quad \text{es} \quad \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} \widehat{C} \\ \widehat{D} \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}.$$

La mejor solución es $\widehat{C} = \frac{9}{7}$, $\widehat{D} = \frac{4}{7}$, y la mejor recta es $\frac{9}{7} + \frac{4}{7}t$.

Observe las hermosas relaciones entre las dos figuras. El problema es el mismo pero el arte lo muestra distinto. En la figura 3.9*b*, b no es una combinación de las columnas $(1, 1, 1)$ y $(-1, 1, 2)$. En la figura 3.9, los tres puntos no están sobre una recta. ¡Los mínimos cuadrados sustituyen los puntos b que no están sobre una recta por puntos p que sí están! Al no ser posible resolver $Ax = b$, se resuelve $A\widehat{x} = p$.

La recta $\frac{9}{7} + \frac{4}{7}t$ tiene alturas $\frac{5}{7}$, $\frac{13}{7}$ y $\frac{17}{7}$ en los instantes de medición −1, 1, 2. ***Estos puntos están sobre una recta***. En consecuencia, el vector $p = \left(\frac{5}{7}, \frac{13}{7}, \frac{17}{7}\right)$ está en el espacio columna. *Este vector es la proyección*. La figura 3.9*b* está en tres dimensiones (o en m dimensiones si hay m puntos) y la figura 3.9*a* está en dos dimensiones (o en n dimensiones si hay n parámetros).

Al restar p de b, los errores son $e = \left(\frac{2}{7}, -\frac{6}{7}, \frac{4}{7}\right)$. Estos son los errores verticales en la figura 3.9*a*, y son las componentes del vector discontinuo de la figura 3.9*b*. Este vector error es ortogonal a la primera columna $(1, 1, 1)$, ya que $\frac{2}{7} - \frac{6}{7} + \frac{4}{7} = 0$. Es ortogonal a la segunda columna $(-1, 1, 2)$ porque $-\frac{2}{7} - \frac{6}{7} + \frac{8}{7} = 0$. Es ortogonal al espacio columna, y está en el espacio nulo izquierdo.

Pregunta: Si las mediciones $b = \left(\frac{2}{7}, -\frac{6}{7}, \frac{4}{7}\right)$ fuesen estos errores, ¿cuál sería la mejor recta y el mejor $\widehat{x}$? Respuesta: La recta cero, que es el eje horizontal, y $\widehat{x} = 0$. Proyección a cero.

Fácilmente pueden resumirse las ecuaciones para ajustar una recta. La primera columna de A contiene a los 1s y la segunda columna a los tiempos t_i. En consecuencia, $A^{\mathrm{T}}A$ contiene a la suma de los 1s y los t_i, así como a t_i^2:

30 Las mediciones $b_1, \ldots, b_m$ están dadas a puntos distintos $t_1, \ldots, t_m$. Así, la recta $\widehat{C} + \widehat{D}t$ que minimiza E^2 proviene de mínimos cuadrados:

$$A^{\mathrm{T}}A \begin{bmatrix} \widehat{C} \\ \widehat{D} \end{bmatrix} = A^{\mathrm{T}}b \quad \text{o bien,} \quad \begin{bmatrix} m & \sum t_i \\ \sum t_i & \sum t_i^2 \end{bmatrix} \begin{bmatrix} \widehat{C} \\ \widehat{D} \end{bmatrix} = \begin{bmatrix} \sum b_i \\ \sum t_i b_i \end{bmatrix}.$$

Observación Las matemáticas de los mínimos cuadrados no están limitadas a ajustar los datos por rectas. En muchos experimentos no hay razón para esperar una relación lineal, y no tendría sentido buscar una. Suponga que se manipula material radiactivo. La salida b es la lectura de un contador Geiger en varios instantes t. Podría saberse que se está trabajando con una mezcla de dos productos químicos, de los cuales se conocen sus vidas medias

(o tasas de desintegración), aunque se ignora la cantidad que se está manipulando de cada producto. Si estas dos cantidades desconocidas son C y D, entonces las lecturas del contador Geiger se comportarían como la suma de dos exponenciales (y no como una recta):

$$b = Ce^{-\lambda t} + De^{-\mu t}. \qquad (8)$$

En la práctica, el contador Geiger no es exacto. En vez de ello, se toman lecturas $b_1, \ldots, b_m$ en los instantes $t_1, \ldots, t_m$, y la ecuación (8) se cumple aproximadamente:

$$Ax = b \quad \text{es} \quad \begin{aligned} Ce^{-\lambda t_1} + De^{-\mu t_1} &\approx b_1 \\ &\vdots \\ Ce^{-\lambda t_m} + De^{-\mu t_m} &\approx b_m. \end{aligned}$$

Si hay más de dos lecturas, $m > 2$, entonces con toda probabilidad no es posible resolver para C y D. Sin embargo, el principio de mínimos cuadrados proporciona valores óptimos $\widehat{C}$ y $\widehat{D}$.

La situación sería completamente diferente si se conocieran las cantidades C y D, y estuviera intentándose descubrir las tasas de desintegración λ y μ. Este es un problema de *mínimos cuadrados no lineales*, y es más difícil. Sin embargo, aún se forman E^2, y la suma de los cuadrados de los errores, y se minimizan. Pero igualar a cero las derivadas no proporciona ecuaciones lineales para las λ y μ óptimas. En los ejercicios sólo se consideran mínimos cuadrados lineales.

Mínimos cuadrados ponderados

Un problema sencillo de mínimos cuadrados es estimar $\widehat{x}$ del peso de un paciente a partir de dos observaciones, $x = b_1$ y $x = b_2$. A menos que $b_1 = b_2$, la situación supone resolver un sistema inconsistente de dos ecuaciones en una incógnita:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} [x] = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

Hasta el momento, b_1 y b_2 se han aceptado como igualmente confiables. Se buscaba el valor $\widehat{x}$ que minimizaba $E^2 = (x - b_1)^2 + (x - b_2)^2$:

$$\frac{dE^2}{dx} = 0 \quad \text{en} \quad \widehat{x} = \frac{b_1 + b_2}{2}.$$

El $\widehat{x}$ óptimo es el promedio. La misma conclusión se obtiene a partir de $A^TA\widehat{x} = A^Tb$. De hecho, A^TA no es una matriz de 1 por 1, y la ecuación normal es $2\widehat{x} = b_1 + b_2$.

Ahora se supone que dos observaciones no son confiables en la misma medida. El valor $x = b_1$ puede obtenerse a partir de una escala más precisa, o bien, en un problema estadístico, de una muestra más grande, que $x = b_2$. A pesar de ello, si b_2 contiene alguna información, no es aconsejable depender totalmente de b_1. El compromiso más sencillo es asignar valores distintos w_1^2 y w_2^2, y escoger $\widehat{x}_W$ que minimice la *suma ponderada de cuadrados*:

$$\textbf{Error ponderado} \quad E^2 = w_1^2(x - b_1)^2 + w_2^2(x - b_2)^2.$$

Si $w_1 > w_2$, entonces a b_1 se asigna mayor importancia. El proceso de minimización (derivada = 0) intenta con mayor ahínco hacer pequeño a $(x - b_1)^2$:

$$\frac{dE^2}{dx} = 2\left[w_1^2(x - b_1) + w_2^2(x - b_2)\right] = 0 \quad \text{en} \quad \widehat{x}_W = \frac{w_1^2 b_1 + w_2^2 b_2}{w_1^2 + w_2^2}. \qquad (9)$$

En vez del promedio de b_1 y b_2 (para $w_1 = w_2 = 1$), $\widehat{x}_W$ es un ***promedio ponderado*** de los datos. Este promedio está más próximo de b_1 que de b_2.

El problema normal de mínimos cuadrados que llevaba a $\widehat{x}_W$ proviene de cambiar $Ax = b$ al nuevo sistema $WAx = Wb$. ***Con esto, la solución cambia de $\widehat{x}$ a $\widehat{x}_W$.*** La matriz W^TW se voltea sobre ambos lados de las ecuaciones normales ponderadas:

La solución por mínimos cuadrados de $WAx = Wb$ es $\widehat{x}_W$:

Ecuaciones normales ponderadas: $(A^TW^TWA)\widehat{x}_W = A^TW^TWb.$

¿Qué ocurre con la representación de b proyectada en $A\widehat{x}$? La proyección $A\widehat{x}_W$ sigue siendo el punto en el espacio columna más próximo a b. Pero la expresión "más próximo" adquiere un nuevo significado cuando la longitud implica a W. La *longitud ponderada* de x es igual a la longitud ordinaria de Wx. Perpendicularidad ya no significa $y^Tx = 0$; en el nuevo sistema la prueba es $(Wy)^T(Wx) = 0$. La matriz W^TW aparece en medio. En este nuevo sentido, la proyección $A\widehat{x}_W$ y el error $b - A\widehat{x}_W$ nuevamente son perpendiculares.

El último párrafo describe ***todos los productos internos***: provienen de matrices invertibles W. Sólo implican a la combinación simétrica $C = W^TW$. ***El producto interno de x y y es y^TCx.*** Para una matriz ortogonal $W = Q$, cuando esta combinación es $C = Q^TQ = I$, el producto interno no es nuevo o diferente. La rotación del espacio deja sin cambio al producto interno. Cualquier otra W cambia la longitud y el producto interno.

Para cualquier matriz invertible W, estas reglas definen un nuevo producto interno y una nueva longitud:

$$\textbf{Ponderado por } W \quad (x, y)_W = (Wy)^T(Wx) \qquad \text{y} \qquad \|x\|_W = \|Wx\|. \qquad (10)$$

Debido a que W es invertible, a ningún vector se asigna la longitud cero (excepto al vector cero). Todos los productos internos posibles —que dependen linealmente de x y y— son positivos cuando $x = y \neq 0$, y se encuentran de esta manera, para alguna matriz $C = W^TW$.

En la práctica, la cuestión importante es la elección de C. La mejor respuesta proviene de los expertos en estadística, originalmente de Gauss. Puede saberse que el error medio es cero. Este es el "valor esperado" del error en b, ¡aunque en realidad se espera que el error no sea cero! También es posible conocer el ***promedio del cuadrado*** del error; ésta es la ***varianza***. Si los errores en las b_i son independientes entre sí, y sus varianzas son σ_i^2, entonces ***los pesos correctos son $w_i = 1/\sigma_i$***. Una medición más precisa, que significa menor varianza, obtiene un mayor peso.

Además de tener confiabilidad desigual, *las observaciones pueden no ser independientes*. Si los errores vienen por parejas –las encuestas para presidente no son independientes de las encuestas para senador, y ciertamente tampoco de las encuestas para vicepresidente— entonces W tiene términos fuera de la diagonal. La mejor matriz no sesgada $C = W^TW$ es la ***inversa de la matriz de covarianza***, cuyo elemento i, j es el valor esperado de (error en b_i) multiplicado por (error en b_j). Entonces, la diagonal principal de C^{-1} contiene a las varianzas σ_i^2, que son el promedio de (error en b_i)^2.

Ejemplo 3 Suponga que dos compañeros en una partida de bridge conjeturan (después de pujar) la cantidad de picas que tienen. Para cada conjetura, los errores –1, 0, 1 pueden tener la misma probabilidad, $\frac{1}{3}$. Así, el error esperado es cero y la varianza es $\frac{2}{3}$:

$$E(e) = \tfrac{1}{3}(-1) + \tfrac{1}{3}(0) + \tfrac{1}{3}(1) = 0$$

$$E(e^2) = \tfrac{1}{3}(-1)^2 + \tfrac{1}{3}(0)^2 + \tfrac{1}{3}(1)^2 = \tfrac{2}{3}.$$

Las dos conjeturas son dependientes, porque están basadas en la misma puja, pero no son idénticas, porque los compañeros tienen manos distintas. Considere que la posibilidad de que los compañeros tengan muchas picas o tengan pocas picas es cero, pero que la posibilidad de errores opuestos es $\frac{1}{3}$. Entonces $E(e_1e_2) = \frac{1}{3}(-1)$, y la inversa de la matriz de covarianza es W^TW:

$$\begin{bmatrix} E(e_1^2) & E(e_1e_2) \\ E(e_1e_2) & E(e_2^2) \end{bmatrix}^{-1} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix}^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = C = W^TW.$$

Esta matriz va en medio de las ecuaciones normales ponderadas.

Conjunto de problemas 3.3

1. Encuentre la mejor solución por mínimos cuadrados $\widehat{x}$ de $3x = 10$, $4x = 5$. ¿Cuál es el error E^2 que se minimiza? Compruebe que el vector error $(10 - 3\widehat{x}, 5 - 4\widehat{x})$ es perpendicular a la columna (3, 4).

2. Suponga que los valores $b_1 = 1$ y $b_2 = 7$ multiplicados por $t_1 = 1$ y $t_2 = 2$ son ajustados por una recta $b = Dt$ *que pasa por el origen*. Resuelva $D = 1$ y $2D = 7$ por mínimos cuadrados. Dibuje la mejor recta.

3. Resuelva $Ax = b$ por mínimos cuadrados, y encuentre $p = A\widehat{x}$ si

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Compruebe que el error $b - p$ es perpendicular a las columnas de A.

4. Escriba $E^2 = \|Ax - b\|^2$ e iguale a cero sus derivadas respecto de u y v, si

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad x = \begin{bmatrix} u \\ v \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}.$$

Compare las ecuaciones resultantes con $A^TA\widehat{x} = A^Tb$, confirmando que el cálculo, así como la geometría, proporcionan las ecuaciones normales. Encuentre la solución $\widehat{x}$ y la proyección $p = A\widehat{x}$. ¿Por qué $p = b$?

5. El siguiente sistema no tiene solución:

$$Ax = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 9 \end{bmatrix} = b.$$

Dibuje y resuelva un ajuste por una recta que lleve a la minimización de la cuadrática $(C - D - 4)^2 + (C - 5)^2 + (C + D - 9)^2$. ¿Cuál es la proyección de b sobre el espacio columna de A?

6. Encuentre la proyección de b sobre el espacio columna de A:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -2 & 4 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 2 \\ 7 \end{bmatrix}.$$

Separe b en $p + q$, con p en el espacio columna y q perpendicular a este espacio. ¿Cuál de los cuatro subespacios contiene a q?

7. Encuentre la matriz proyección P sobre el espacio generado por $a_1 = (1, 0, 1)$ y $a_2 = (1, 1, -1)$.

8. Si P es la matriz proyección sobre un subespacio k-dimensional $\mathbf{S}$ de todo el espacio $\mathbf{R}^n$, ¿cuál es el espacio columna de P, y cuál es su rango?

9. *a*) Si $P = P^TP$, demuestre que P es una matriz proyección.
b) ¿Sobre qué subespacio proyecta la matriz $P = 0$?

10. Si los vectores a_1, a_2, y b son ortogonales, ¿cuáles son A^TA y A^Tb? ¿Cuál es la proyección de b sobre el plano de a_1 y a_2?

11. Suponga que P es la matriz proyección sobre el subespacio $\mathbf{S}$ y que Q es la matriz proyección sobre los complementos ortogonales $\mathbf{S}^{\perp}$. ¿Cuáles son $P + Q$ y PQ? Demuestre que $P - Q$ es su propia inversa.

12. Si $\mathbf{V}$ es el subespacio generado por (1, 1, 0, 1) y (0, 0, 1, 0), encuentre
a) Una base para el complemento ortogonal $\mathbf{V}^{\perp}$.
b) La matriz proyección P sobre $\mathbf{V}$.
c) El vector en $\mathbf{V}$ más próximo al vector $b = (0, 1, 0, -1)$ en $\mathbf{V}^{\perp}$.

13. Encuentre el mejor ajuste por una recta (mínimos cuadrados) de las mediciones

$$\begin{array}{llll} b = 4 \text{ en } t = -2, & b = 3 \text{ en } t = -1, \\ b = 1 \text{ en } t = 0, & b = 0 \text{ en } t = 2. \end{array}$$

Luego encuentre la proyección de $b = (4, 3, 1, 0)$ sobre el espacio columna de

$$A = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 2 \end{bmatrix}.$$

14. Los vectores $a_1 = (1, 1, 0)$ y $a_2 = (1, 1, 1)$ generan un plano en $\mathbf{R}^3$. Encuentre la matriz proyección P sobre el plano, y encuentre un vector b diferente de cero que sea proyectado en cero.

15. Si P es la matriz proyección sobre una recta en el plano x-y, dibuje una figura para describir el efecto de la "matriz reflexión" $H = I - 2P$. Explique geométrica y algebraicamente por qué $H^2 = I$.

16. Demuestre que si la longitud de u es uno, entonces la matriz $P = uu^T$ de rango 1 es una matriz proyección: tiene las propiedades i) y ii) enunciadas en 3N. Si se escoge $u = a/\|a\|$, P se convierte en la matriz proyección sobre la recta que pasa por a, y Pb es el punto $p = \widehat{x}a$. Las proyecciones de rango 1 corresponden exactamente a problemas de mínimos cuadrados en una incógnita.

17. ¿Qué matriz de 2 por 2 proyecta el plano x-y sobre la recta de 45° $x + y = 0$?

18. Se quiere ajustar un plano $y = C + Dt + Ez$ para los cuatro puntos

$$\begin{array}{ll} y = 3 \text{ en } t = 1,\ z = 1 & y = 6 \text{ en } t = 0,\ z = 3 \\ y = 5 \text{ en } t = 2,\ z = 1 & y = 0 \text{ en } t = 0,\ z = 0. \end{array}$$

a) Encuentre 4 ecuaciones en 3 incógnitas para hacer pasar un plano por los puntos (en caso de que tal plano exista).
b) Encuentre 3 ecuaciones en 3 incógnitas para la mejor solución por mínimos cuadrados.

19. Si $P_C = A(A^TA)^{-1}A^T$ es la proyección sobre el espacio columna de A, ¿cuál es la proyección P_R sobre el espacio renglón? (¡No es P_C^T!)

20. Si P es la proyección sobre el espacio columna de A, ¿cuál es la proyección sobre el espacio nulo izquierdo?

21. Suponga que L_1 es la recta que pasa por el origen en la dirección de a_1 y que L_2 es la recta que pasa por b en la dirección de a_2. Para encontrar los puntos más próximos x_1a_1 y $b + x_2a_2$ sobre las dos rectas, escriba las dos ecuaciones para x_1 y x_2 que minimizan $\|x_1a_1 - x_2a_2 - b\|$. Resuelva para x si $a_1 = (1, 1, 0)$, $a_2 = (0, 1, 0)$, $b = (2, 1, 4)$.

22. Encuentre la mejor recta $C + Dt$ para ajustar $b = 4, 2, -1, 0, 0$ en los instantes $t = -2, -1, 0, 1, 2$.

23. Demuestre que el mejor ajuste por mínimos cuadrados a un conjunto de mediciones $y_1, \ldots, y_m$ por una *recta horizontal* (una función constante $y = C$) es su promedio

$$C = \frac{y_1 + \cdots + y_m}{m}.$$

24. Encuentre el mejor ajuste por una recta a las siguientes mediciones, y dibuje su solución:

$$y = 2 \text{ en } t = -1, \quad y = 0 \text{ en } t = 0,$$
$$y = -3 \text{ en } t = 1, \quad y = -5 \text{ en } t = 2.$$

25. Suponga que en vez de, por medio de una recta, los datos del problema 24 se ajustan por una parábola: $y = C + Dt + Et^2$. En el sistema inconsistente $Ax = b$ proveniente de las cuatro mediciones, ¿cuáles son la matriz de coeficientes A, el vector desconocido x, y el vector de datos b? No es necesario calcular $\widehat{x}$.

26. Bajo tortura, una persona de la Edad Media fue alargada las longitudes $L = 5, 6$, y 7 pies bajo fuerzas aplicadas de $F = 1, 2$, y 4 toneladas. Suponga la ley de Hooke $L = a + bF$ para encontrar por mínimos cuadrados la longitud normal a de la persona.

En los problemas 27 a 31 se presentan los conceptos básicos de la estadística: la base de los mínimos cuadrados.

27. (Recomendado) Este problema proyecta $b = (b_1, \ldots, b_m)$ sobre la recta que pasa por $a = (1, \ldots, 1)$. Se resuelven m ecuaciones $ax = b$ en 1 incógnita (por mínimos cuadrados).
a) Resuelva $a^Ta\widehat{x} = a^Tb$ para demostrar que $\widehat{x}$ es *la media* (el promedio) de las bs.
b) Encuentre $e = b - a\widehat{x}$, la *varianza* $\|e\|^2$, y la *desviación estándar* $\|e\|$.
c) La recta horizontal $\widehat{b} = 3$ es la más próxima a $b = (1, 2, 6)$. Compruebe que $p = (3, 3, 3)$ es perpendicular a e, y encuentre la matriz proyección P.

28. Primera suposición detrás de los mínimos cuadrados: La ***media*** de cada error de medición ***es cero***. Multiplique los 8 vectores error $b - Ax = (\pm 1, \pm 1, \pm 1)$ por $(A^TA)^{-1}A^T$ para demostrar que el promedio de los 8 vectores $\widehat{x} - x$ también es cero. La estimación $\widehat{x}$ *es insesgada*.

29. Segunda suposición detrás de los mínimos cuadrados: Los m errores e_i son independientes de la varianza σ^2, de modo que el promedio $(b - Ax)(b - Ax)^T$ es $\sigma^2 I$. Multiplique por la izquierda por $(A^TA)^{-1}A^T$ y por la derecha por $A(A^TA)^{-1}$ para demostrar que el promedio de $(\widehat{x} - x)(\widehat{x} - x)^T$ es $\sigma^2(A^TA)^{-1}$. Esta es la fundamental ***matriz de covarianza*** para el error en $\widehat{x}$.

30. Un doctor toma 4 lecturas del ritmo cardiaco de una persona. La mejor solución de $x = b_1, \ldots, x = b_4$ es el promedio $\widehat{x}$ de $b_1, \ldots, b_4$. La matriz A es la columna de 1s. En el problema 29 se proporciona el error esperado $(\widehat{x} - x)^2$ como $\sigma^2(A^TA)^{-1} =$ ____. Al promediar, la varianza cae desde σ^2 hasta $\sigma^2/4$.

31. Si se conoce el promedio $\widehat{x}_9$ de 9 números $b_1, \ldots, b_9$, ¿cuán rápido puede encontrarse el promedio $\widehat{x}_{10}$ con un número más b_{10}? La idea de mínimos cuadrados *recurrentes* es evitar la suma de 10 números. ¿Qué coeficiente de $\widehat{x}_9$ proporciona correctamente $\widehat{x}_{10}$?

$$\widehat{x}_{10} = \tfrac{1}{10}b_{10} + \underline{\quad}\,\widehat{x}_9 = \tfrac{1}{10}(b_1 + \cdots + b_{10}).$$

En los problemas 32 a 37, use cuatro puntos $b = (0, 8, 8, 20)$ para producir ideas nuevas.

32. Con $b = 0, 8, 8, 20$ en $t = 0, 1, 3, 4$, escriba y resuelva las ecuaciones normales $A^TA\widehat{x} = A^Tb$. Para la mejor recta como en la figura 3.9*a*, encuentre sus cuatro alturas p_i y sus cuatro errores e_i. ¿Cuál es el valor mínimo $E^2 = e_1^2 + e_2^2 + e_3^2 + e_4^2$?

33. (La recta $C + Dt$ pasa por los ps). Con $b = 0, 8, 8, 20$ en los tiempos $t = 0, 1, 3, 4$, escriba las cuatro ecuaciones $Ax = b$ (irresoluble). Cambie las mediciones a $p = 1, 5, 13, 17$ y encuentre una solución exacta de $A\widehat{x} = p$.

34. Compruebe que $e = b - p = (-1, 3, -5, 3)$ es perpendicular a ambas columnas de A. ¿Cuál es la distancia más corta $\|e\|$ de b al espacio columna de A?

35. Para la parábola más próxima $b = C + Dt + Et^2$ a los mismos cuatro puntos, escriba las ecuaciones irresolubles $Ax = b$ en tres incógnitas $x = (C, D, E)$. Escriba las tres ecuaciones normales $A^TA\widehat{x} = A^Tb$ (no se requiere su solución). Ahora se está ajustando una parábola a los cuatro puntos; ¿qué ocurre en la figura 3.9*b*?

36. Para la cúbica más próxima $b = C + Dt + Et^2 + Ft^3$ a los mismos cuatro puntos, escriba las cuatro ecuaciones $Ax = b$. Resuélvalas por eliminación. Ahora, esta cúbica pasa exactamente por los puntos. ¿Cuáles son p y e?

37. El promedio de los cuatro tiempos es $\widehat{t} = \frac{1}{4}(0 + 1 + 3 + 4) = 2$. El promedio de las cuatro bs es $\widehat{b} = \frac{1}{4}(0 + 8 + 8 + 20) = 9$.
a) Compruebe que la mejor recta *pasa por el punto central* $(\widehat{t}, \widehat{b}) = (2, 9)$.
b) Explique por qué $C + D\widehat{t} = \widehat{b}$ proviene de la primera ecuación en $A^TA\widehat{x} = A^Tb$.

38. ¿Qué ocurre con el peso ponderado $\widehat{x}_W = (w_1^2b_1 + w_2^2b_2)/(w_1^2 + w_2^2)$ si el primer peso w_1 tiende a cero? La medición b_1 es completamente desconfiable.

39. A partir de m mediciones independientes $b_1, \ldots, b_m$ del pulso de una persona, ponderadas por $w_1, \ldots, w_m$, ¿cuál es el promedio ponderado que sustituye a la ecuación (9)? La respuesta es la mejor estimación cuando las varianzas estadísticas son $\sigma_i^2 = 1/w_i^2$.

40. Si $W = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$, encuentre el producto interno de $x = (2, 3)$ y $y = (1, 1)$, así como la longitud W de x. ¿Qué recta de vectores es W-perpendicular a y?

41. Encuentre la solución ponderada por mínimos cuadrados $\widehat{x}_W$ de $Ax = b$:

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \qquad b = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \qquad W = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Compruebe que la proyección $A\widehat{x}_W$ sigue siendo perpendicular (¡en el producto interno W!) al error $b - A\widehat{x}_W$.

42. *a)* Suponga que usted adivina la edad de su profesor, cometiendo errores $e = -2, -1, 5$, con probabilidades $\frac{1}{2}, \frac{1}{4}, \frac{1}{4}$. Compruebe que el error esperado $E(e)$ es cero, y encuentre la varianza $E(e^2)$.

b) Si también el profesor adivina (o intenta recordar), cometiendo errores $-1, 0, 1$, con probabilidades $\frac{1}{8}, \frac{6}{8}, \frac{1}{8}$, ¿qué pesos w_1 y w_2 proporcionan la confiabilidad de su conjetura, así como la confiabilidad de la conjetura del profesor?

3.4 BASES ORTOGONALES Y GRAM-SCHMIDT

En una base ortogonal, todos los vectores son perpendiculares entre sí. Los ejes de coordenadas son mutuamente ortogonales. Esta situación es casi óptima, y la única mejoría es fácil de realizar: cada vector se divide entre su longitud con la finalidad de hacerlo un *vector unitario*. Así se cambia una base ***ortogonal*** a una base ***ortonormal*** de *q*s:

3P Los vectores $q_1, \ldots, q_n$ son ***ortonormales*** si

$$q_i^{\mathrm{T}} q_j = \begin{cases} 0 \text{ siempre que } i \neq j, & \text{proporcionando la ortogonalidad;} \\ 1 \text{ siempre que } i = j, & \text{proporcionando la normalización.} \end{cases}$$

Una matriz con columnas ortonormales se denomina Q.

El ejemplo más importante es la *base estándar*. Para el plano *x-y*, los ejes más conocidos $e_1 = (1, 0)$ y $e_2 = (0, 1)$ no sólo son perpendiculares, sino que también son horizontal y vertical, respectivamente. Q es la matriz identidad de 2 por 2. En n dimensiones, la base estándar $e_1, \ldots, e_n$ nuevamente consta de *las columnas de* $Q = I$:

Base estándar
$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad \ldots, \qquad e_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

¡Esta no es la única base ortonormal! Es posible rotar los ejes sin modificar los ángulos rectos a los que se cortan. Estas matrices rotación constituyen ejemplos de Q.

Si se tiene un subespacio de $\mathbf{R}^n$, los vectores estándar e_i pueden no estar en ese subespacio. Sin embargo, el subespacio siempre tiene una base ortonormal, que puede construirse en forma sencilla a partir de cualquier base dada. Esta construcción, que transforma un conjunto sesgado de ejes en un conjunto perpendicular, se denomina ***ortogonalización de Gram-Schmidt***.

En resumen, los tres tópicos fundamentales de esta sección son:

1. La definición y las propiedades de las matrices ortogonales Q.
2. La solución de $Qx = b$, ya sea de n por n o rectangular (mínimos cuadrados).
3. El proceso de Gram-Schmidt y su interpretación como una nueva factorización $A = QR$.

Matrices ortogonales

3Q Si Q (cuadrada o rectangular) tiene columnas ortonormales, entonces $Q^{\mathrm{T}}Q = I$:

Columnas ortonormales
$$\begin{bmatrix} — q_1^{\mathrm{T}} — \\ — q_2^{\mathrm{T}} — \\ \vdots \\ — q_n^{\mathrm{T}} — \end{bmatrix} \begin{bmatrix} | & | & & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdot & 0 \\ 0 & 1 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & 1 \end{bmatrix} = I. \quad (1)$$

Una matriz ortogonal es una matriz cuadrada con columnas ortonormales.* Por tanto, Q^{T} es Q^{-1}. Para matrices cuadradas ortogonales, ***la traspuesta es la inversa.***

Cuando el renglón i de Q^{T} multiplica a la columna j de Q, el resultado es $q_i^{\mathrm{T}}q_j = 0$. Sobre la diagonal, donde $i = j$, se tiene $q_i^{\mathrm{T}}q_i = 1$. Esta es la normalización a vectores unitarios de longitud 1.

Observe que $Q^{\mathrm{T}}Q = I$ incluso si Q es rectangular. Pero entonces Q^{T} es sólo una inversa izquierda.

Ejemplo 1

$$Q = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}, \qquad Q^{\mathrm{T}} = Q^{-1} = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}.$$

Q hace girar a todo vector un ángulo θ, y Q^{T} lo hace girar de regreso un ángulo $-\theta$. Resulta evidente que las columnas son ortogonales, y son ortonormales porque $\operatorname{sen}^2\theta + \cos^2\theta = 1$. La matriz Q^{T} es justo una matriz tan ortogonal como lo es Q.

Ejemplo 2 Cualquier matriz permutación P es una matriz ortogonal. Ciertamente, las columnas son vectores unitarios y ortogonales, ya que el 1 aparece en un sitio distinto en cada columna: La traspuesta es la inversa.

$$\text{Si } P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \text{ entonces } P^{-1} = P^{\mathrm{T}} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Una P antidiagonal, con $P_{13} = P_{22} = P_{31} = 1$, lleva los ejes *x-y-z* en los ejes *z-y-x*: un sistema "derecho" en un sistema "izquierdo". Por tanto, hubiera sido erróneo sugerir que cada Q ortogonal representa una rotación. *También se permite una reflexión.* $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ refleja cada punto (x, y) en (y, x), su imagen especular a través de la recta a 45°. Geométricamente, una Q ortogonal es el producto de una rotación y una reflexión.

Queda por abordar una propiedad compartida por las rotaciones y las reflexiones, y de hecho, por cualquier matriz ortogonal. No es compartida por las proyecciones, que no son ortogonales y ni siquiera invertibles. Las proyecciones reducen la longitud de un vector, mientras que las matrices ortogonales poseen una propiedad que es la más importante y característica de todas:

*Matriz ortonormal hubiera sido un mejor nombre, pero ya es demasiado tarde para cambiar. También, no hay una denominación aceptada para una matriz rectangular con columnas ortonormales. Seguimos escribiendo Q, pero no se denominará "matriz ortogonal" a menos que sea cuadrada.

3R La multiplicación por cualquier Q preserva las longitudes:

Longitudes sin cambio $\|Qx\| = \|x\|$ para todo vector x. (2)

También preserva productos internos y ángulos, ya que $(Qx)^{\mathrm{T}}(Qy) = x^{\mathrm{T}}Q^{\mathrm{T}}Qy = x^{\mathrm{T}}y$.

La preservación de longitudes proviene directamente de $Q^{\mathrm{T}}Q = I$:

$$\|Qx\|^2 = \|x\|^2 \quad \text{debido a que} \quad (Qx)^{\mathrm{T}}(Qx) = x^{\mathrm{T}}Q^{\mathrm{T}}Qx = x^{\mathrm{T}}x. \tag{3}$$

Cuando el espacio se gira o refleja, se preservan todos los productos internos y las longitudes.

A continuación se abordará el cálculo en que se utiliza la propiedad especial $Q^{\mathrm{T}} = Q^{-1}$. Si se tiene una base, entonces cualquier vector es una combinación de los vectores de la base. Esto es excepcionalmente simple para una base ortonormal, lo cual será un concepto clave detrás de las series de Fourier. El problema *es encontrar los coeficientes de los vectores de la base*:

b se escribe como una combinación $b = x_1q_1 + x_2q_2 + \cdots + x_nq_n$.

Para calcular x_1 hay un truco evidente. *Ambos lados de la ecuación se multiplican por* q_1^{T}. En el miembro izquierdo está $q_1^{\mathrm{T}}b$. En el miembro derecho todos los términos desaparecen (porque $q_1^{\mathrm{T}}q_j = 0$) excepto el primer término. Lo que queda es

$$q_1^{\mathrm{T}}b = x_1q_1^{\mathrm{T}}q_1.$$

Debido a que $q_1^{\mathrm{T}}q_1 = 1$, se ha encontrado que $x_1 = q_1^{\mathrm{T}}b$. De manera semejante, el segundo coeficiente es $x_2 = q_2^{\mathrm{T}}b$; este término permanece cuando se multiplica por q_2^{T}. Los otros términos desaparecen por la ortogonalidad. Cada pieza de b tiene una fórmula sencilla, y al recombinar las piezas se recupera b:

Todo vector b es igual a $(q_1^{\mathrm{T}}b)q_1 + (q_2^{\mathrm{T}}b)q_2 + \cdots + (q_n^{\mathrm{T}}b)q_n$. (4)

No puedo resistir escribir esta base ortonormal en una matriz cuadrada Q. La ecuación vectorial $x_1q_1 + \cdots + x_nq_n = b$ es idéntica a $Qx = b$. (Las columnas de Q multiplican a las componentes de x). Su solución es $x = Q^{-1}b$. Pero como $Q^{-1} = Q^{\mathrm{T}}$ –aquí es donde entra la ortonormalidad– la solución también es $x = Q^{\mathrm{T}}b$:

$$x = Q^{\mathrm{T}}b = \begin{bmatrix} \text{—— } q_1^{\mathrm{T}} \text{ ——} \\ \\ \text{—— } q_n^{\mathrm{T}} \text{ ——} \end{bmatrix} \begin{bmatrix} \\ b \\ \\ \end{bmatrix} = \begin{bmatrix} q_1^{\mathrm{T}}b \\ \\ q_n^{\mathrm{T}}b \end{bmatrix} \tag{5}$$

Las componentes de x son los productos internos $q_i^{\mathrm{T}}b$, como en la ecuación (4).

La forma matricial también muestra lo que ocurre cuando las columnas *no son* ortonormales. Expresar b como una combinación $x_1a_1 + \cdots + x_na_n$ es lo mismo que resolver $Ax = b$. Los vectores de la base se van en las columnas de A. En ese caso, se requiere A^{-1}, lo cual cuesta trabajo. En el caso ortonormal sólo se requiere Q^{T}.

Observación 1 La razón $a^{\mathrm{T}}b/a^{\mathrm{T}}a$ ya apareció antes, cuando b se proyectó sobre una recta. Aquí a es q_1, el denominador es 1 y la proyección es $(q_1^{\mathrm{T}}b)q_1$. Por tanto, se tiene una nueva interpretación de la fórmula (4): *Todo vector b es la suma de sus proyecciones unidimensionales sobre las rectas que pasan por las qs.*

Debido a que estas proyecciones son ortogonales, el teorema de Pitágoras sigue siendo válido. El cuadrado de la hipotenusa sigue siendo la suma de los cuadrados de las componentes:

$$\|b\|^2 = (q_1^{\mathrm{T}}b)^2 + (q_2^{\mathrm{T}}b)^2 + \cdots + (q_n^{\mathrm{T}}b)^2 \quad \text{que es} \quad \|Q^{\mathrm{T}}b\|^2. \tag{6}$$

Observación 2 Debido a que $Q^{\mathrm{T}} = Q^{-1}$, también se tiene que $QQ^{\mathrm{T}} = I$. Cuando Q viene antes que Q^{T}, la multiplicación toma los productos internos de los *renglones* de Q. (Para $Q^{\mathrm{T}}Q$ eran las columnas.) Ya que el resultado es nuevamente la matriz identidad, se llega a una conclusión sorprendente: ***Los renglones de una matriz cuadrada son ortonormales siempre que lo sean las columnas.*** Los renglones apuntan en direcciones completamente distintas a las de las columnas, y geométricamente no veo por qué están obligados a ser ortonormales, pero lo son.

Columnas ortonormales
Renglones ortonormales

$$Q = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix}.$$

Matrices rectangulares con columnas ortonormales

Este capítulo es sobre $Ax = b$, cuando A no necesariamente es cuadrada. Para $Qx = b$ ahora se admite la misma posibilidad: puede haber más renglones que columnas. Los n vectores ortonormales q_i en las columnas de Q tienen $m > n$ componentes. Así, Q es una matriz de m por n y no es de esperar poder resolver $Qx = b$ de manera exacta. *Se resuelve por mínimos cuadrados.*

Si hay justicia, las columnas ortonormales deben hacer más sencillo el problema. Funcionó para matrices cuadradas, y ahora funcionará para matrices rectangulares. La clave consiste en observar que *se sigue teniendo* $Q^{\mathrm{T}}Q = I$. Así, Q^{T} sigue siendo la ***inversa izquierda*** de Q.

Para mínimos cuadrados, esto es todo lo que se necesita. Las ecuaciones normales se presentaron al multiplicar $Ax = b$ por la matriz traspuesta, para obtener $A^{\mathrm{T}}A\widehat{x} = A^{\mathrm{T}}b$. Ahora las ecuaciones normales son $Q^{\mathrm{T}}Q\widehat{x} = Q^{\mathrm{T}}b$. ¡Pero $Q^{\mathrm{T}}Q$ es la matriz identidad! En consecuencia, $\widehat{x} = Q^{\mathrm{T}}b$, ya sea que Q es cuadrada y si $\widehat{x}$ es una solución exacta, o si Q es rectangular y se requieren mínimos cuadrados.

3S Si las columnas de Q son ortonormales, el problema de mínimos cuadrados se facilita:

$Qx = b$	sistema rectangular sin solución para casi todo b.
$Q^{\mathrm{T}}Q\widehat{x} = Q^{\mathrm{T}}b$	ecuación normal para el mejor $\widehat{x}$ donde $Q^{\mathrm{T}}Q = I$.
$\widehat{x} = Q^{\mathrm{T}}b$	$\widehat{x}_i$ es $q_i^{\mathrm{T}}b$.
$p = Q\widehat{x}$	la proyección de b es $(q_1^{\mathrm{T}}b)q_1 + \cdots + (q_n^{\mathrm{T}}b)q_n$.
$p = QQ^{\mathrm{T}}b$	la matriz proyección es $P = QQ^{\mathrm{T}}$.

Las últimas fórmulas son como $p = A\widehat{x}$ y $P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$. Cuando las columnas son ortonormales, la "matriz producto cruz" $A^{\mathrm{T}}A$ se convierte en $Q^{\mathrm{T}}Q = I$. La parte difícil de los mínimos cuadrados desaparece cuando los vectores son ortonormales. Las proyecciones sobre los ejes no se presentan por pares, y p es la suma $p = (q_1^{\mathrm{T}}b)q_1 + \cdots + (q_n^{\mathrm{T}}b)q_n$.

Se recalca que estas proyecciones no reconstruyen b. En el caso cuadrado $m = n$ sí lo hacían. En el caso rectangular $m > n$ no lo hacen. Proporcionan la proyección p y no el vector original b, que es todo lo que puede esperarse cuando hay más ecuaciones que incógnitas, y las qs dejan de ser una base. La matriz proyección suele ser $A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$,

y aquí se simplifica a

$$P = Q(Q^{\mathrm{T}}Q)^{-1}Q^{\mathrm{T}} \quad \text{o bien,} \quad P = QQ^{\mathrm{T}}. \tag{7}$$

Observe que $Q^{\mathrm{T}}Q$ es la matriz identidad de n por n, mientras QQ^{T} es una matriz proyección P de m por m. Es la matriz identidad sobre las columnas de Q (P las deja solas). Pero QQ^{T} es la matriz cero sobre el complemento ortogonal (el espacio nulo de Q^{T}).

Ejemplo 3 El siguiente caso es sencillo pero típico. Suponga que un punto $b = (x, y, z)$ se proyecta sobre el plano x-y. Su proyección es $p = (x, y, 0)$, y esta es la suma de las proyecciones por separado sobre los ejes x y y:

$$q_1 = \begin{bmatrix}1\\0\\0\end{bmatrix} \quad \text{y} \quad (q_1^{\mathrm{T}}b)q_1 = \begin{bmatrix}x\\0\\0\end{bmatrix}; \qquad q_2 = \begin{bmatrix}0\\1\\0\end{bmatrix} \quad \text{y} \quad (q_2^{\mathrm{T}}b)q_2 = \begin{bmatrix}0\\y\\0\end{bmatrix}.$$

La matriz proyección global es

$$P = q_1q_1^{\mathrm{T}} + q_2q_2^{\mathrm{T}} = \begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}, \quad \text{y} \quad P\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}x\\y\\0\end{bmatrix}.$$

Proyección sobre un plano = suma de las proyecciones sobre q_1 y q_2 ortonormales.

Ejemplo 4 Cuando el promedio de los tiempos de medición es cero, ajustar una recta conduce a columnas ortogonales. Considere $t_1 = -3$, $t_2 = 0$, $t_3 = 3$. Luego, el intento de ajustar $y = C + Dt$ produce tres ecuaciones en dos incógnitas:

$$\begin{aligned} C + Dt_1 &= y_1 \\ C + Dt_2 &= y_2, \\ C + Dt_3 &= y_3 \end{aligned} \quad \text{o bien,} \quad \begin{bmatrix}1&-3\\1&0\\1&3\end{bmatrix}\begin{bmatrix}C\\D\end{bmatrix} = \begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}.$$

Las columnas $(1, 1, 1)$ *y* $(-3, 0, 3)$ *son ortogonales*. Es posible proyectar y por separado sobre cada columna, de modo que los mejores coeficientes $\widehat{C}$ y $\widehat{D}$ pueden encontrarse por separado:

$$\widehat{C} = \frac{\begin{bmatrix}1&1&1\end{bmatrix}\begin{bmatrix}y_1&y_2&y_3\end{bmatrix}^{\mathrm{T}}}{1^2+1^2+1^2}, \qquad \widehat{D} = \frac{\begin{bmatrix}-3&0&3\end{bmatrix}\begin{bmatrix}y_1&y_2&y_3\end{bmatrix}^{\mathrm{T}}}{(-3)^2+0^2+3^2}.$$

Observe que $\widehat{C} = (y_1 + y_2 + y_3)/3$ es la *media* de los datos. $\widehat{C}$ proporciona el mejor ajuste por una recta horizontal, mientras $\widehat{D}t$ es el mejor ajuste por una recta que pasa por el origen. *Las columnas son ortogonales, de modo que la suma de estas dos piezas por separado constituye el mejor ajuste por cualquier recta.* Las columnas no son vectores unitarios, por lo que $\widehat{C}$ y $\widehat{D}$ tienen la longitud al cuadrado en el denominador.

Las columnas ortogonales son en mucho mejores, tanto que merece la pena cambiar a ese caso. Si el promedio de los tiempos de observación no es cero –es $\bar{t} = (t_1 + \cdots + t_m)/m$– entonces el origen del tiempo puede cambiarse por $\bar{t}$. En vez de $y = C + Dt$ se trabaja con $y = c + d(t - \bar{t})$. ¡La mejor recta es la misma! Así como en el ejemplo, se encuentra

$$\widehat{c} = \frac{\begin{bmatrix}1&\cdots&1\end{bmatrix}\begin{bmatrix}y_1&\cdots&y_m\end{bmatrix}^{\mathrm{T}}}{1^2+1^2+\cdots+1^2} = \frac{y_1+\cdots+y_m}{m}$$

$$\widehat{d} = \frac{\begin{bmatrix}(t_1-\bar{t})&\cdots&(t_m-\bar{t})\end{bmatrix}\begin{bmatrix}y_1&\cdots&y_m\end{bmatrix}^{\mathrm{T}}}{(t_1-\bar{t})^2+\cdots+(t_m-\bar{t})^2} = \frac{\sum(t_i-\bar{t})y_i}{\sum(t_i-\bar{t})^2}. \tag{8}$$

El mejor $\widehat{c}$ es la media, y también se obtiene una fórmula conveniente para $\widehat{d}$. Los elementos fuera de la diagonal de la $A^{\mathrm{T}}A$ anterior eran Σt_i, y al cambiar el tiempo por $\bar{t}$ estos elementos se hacen cero. Este cambio es un ejemplo del proceso de Gram-Schmidt, ***que ortogonaliza la situación de antemano.***

Las matrices ortogonales son cruciales en álgebra lineal numérica, ya que no introducen inestabilidad. Mientras las longitudes permanecen igual, el redondeo está bajo control. La ortogonalización de vectores se ha convertido en una técnica esencial. Quizá sólo esté en segundo lugar, detrás de la eliminación; y conduce a una factorización $A = QR$ que es casi tan famosa como $A = LU$.

El proceso de Gram-Schmidt

Suponga que se tienen tres vectores independientes a, b, c. Si son ortonormales, las cosas son fáciles. Para proyectar un vector v sobre el primero, se calcula $(a^{\mathrm{T}}v)a$. Para proyectar el mismo vector v sobre el plano de los dos primeros, simplemente se suma $(a^{\mathrm{T}}v)a + (b^{\mathrm{T}}v)b$. Para proyectar sobre lo generado por a, b, c, se suman tres proyecciones. Todos los cálculos sólo requieren los productos internos $a^{\mathrm{T}}v$, $b^{\mathrm{T}}v$, y $c^{\mathrm{T}}v$. Pero para hacer realidad esto, es obligado decir: "***Si*** son ortonormales." A continuación se propone encontrar una manera de ***hacerlos*** ortonormales.

El método es sencillo. Se tienen a, b, c y se buscan q_1, q_2, q_3. Con q_1 no hay problema: puede ir en la dirección de a. Se divide entre la longitud, de modo que $q_1 = a/\|a\|$ es un vector unitario. El verdadero problema empieza con q_2, que debe ser ortogonal a q_1. Si el segundo vector b tiene cualquier componente en la dirección de q_1 (que es la dirección de a), ***es necesario restar esa componente***:

$$\textbf{Segundo vector} \qquad B = b - (q_1^{\mathrm{T}}b)q_1 \quad \text{y} \quad q_2 = B/\|B\|. \tag{9}$$

B es ortogonal a q_1. Es la parte de b que va en una nueva dirección, y no en la dirección de a. En la figura 3.10, B es perpendicular a q_1. Establece la dirección para q_2.

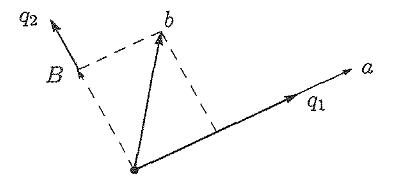


Figura 3.10 La componente q_1 de b se elimina; a y B se normalizan en q_1 y q_2.

En este punto ya se cuenta con q_1 y q_2. La tercera dirección ortogonal empieza con c. No está en el plano de q_1 y q_2, que es el plano de a y b. No obstante, puede tener una componente en ese plano, y ésta, debe restarse. (Si el resultado es $C = 0$, esto indica que a, b,

c no eran independientes desde el principio.) Lo que queda es la componente C que se busca, la parte que está en una nueva dirección perpendicular al plano:

Tercer vector $$C = c - (q_1^{\mathrm{T}}c)q_1 - (q_2^{\mathrm{T}}c)q_2 \quad \text{y} \quad q_3 = C/\|C\|. \qquad (10)$$

Esta es la idea de todo el proceso de Gram-Schmidt; ***restar de cada vector nuevo sus componentes en las direcciones ya establecidas***. Esta idea se utiliza una y otra vez.* Cuando se tiene un cuarto vector, se restan sus componentes en las direcciones de q_1, q_2, q_3.

Ejemplo 5 **Gram-Schmidt** Suponga que los vectores independientes son a, b, c:

$$a = \begin{bmatrix}1\\0\\1\end{bmatrix}, \quad b = \begin{bmatrix}1\\0\\0\end{bmatrix}, \quad c = \begin{bmatrix}2\\1\\0\end{bmatrix}.$$

Para encontrar q_1, el primer vector se hace unitario: $q_1 = a/\sqrt{2}$. Para encontrar q_2, del segundo vector se resta su componente en la primera dirección:

$$B = b - (q_1^{\mathrm{T}}b)q_1 = \begin{bmatrix}1\\0\\0\end{bmatrix} - \frac{1}{\sqrt{2}}\begin{bmatrix}1/\sqrt{2}\\0\\1/\sqrt{2}\end{bmatrix} = \frac{1}{2}\begin{bmatrix}1\\0\\-1\end{bmatrix}.$$

El q_2 normalizado es B dividido entre su longitud, para producir un vector unitario:

$$q_2 = \begin{bmatrix}1/\sqrt{2}\\0\\-1/\sqrt{2}\end{bmatrix}.$$

Para encontrar q_3, de c se resta su componente a lo largo de q_1 y q_2:

$$\begin{aligned} C &= c - (q_1^{\mathrm{T}}c)q_1 - (q_2^{\mathrm{T}}c)q_2 \\ &= \begin{bmatrix}2\\1\\0\end{bmatrix} - \sqrt{2}\begin{bmatrix}1/\sqrt{2}\\0\\1/\sqrt{2}\end{bmatrix} - \sqrt{2}\begin{bmatrix}1/\sqrt{2}\\0\\-1/\sqrt{2}\end{bmatrix} = \begin{bmatrix}0\\1\\0\end{bmatrix}. \end{aligned}$$

Este vector ya es unitario, de modo que es q_3. Acudí a longitudes desesperadas para disminuir el número de raíces cuadradas (la parte penosa del proceso de Gram-Schmidt). El resultado es un conjunto de vectores ortonormales q_1, q_2, q_3, que van en las columnas de una matriz ortonormal Q:

Base ortonormal $$Q = \begin{bmatrix} & & \\ q_1 & q_2 & q_3 \\ & & \end{bmatrix} = \begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2} & 0\\ 0 & 0 & 1\\ 1/\sqrt{2} & -1/\sqrt{2} & 0\end{bmatrix}.$$

*Si Gram fue el primero que tuvo esta idea, ¿qué le quedó a Schmidt?

3T El proceso de Gram-Schmidt empieza con vectores independientes $a_1, \ldots, a_n$ y termina con vectores ortonormales $q_1, \ldots, q_n$. En el paso j resta de a_j sus componentes en las direcciones $q_1, \ldots, q_{j-1}$ que ya están establecidas:

$$A_j = a_j - (q_1^{\mathrm{T}}a_j)q_1 - \cdots - (q_{j-1}^{\mathrm{T}}a_j)q_{j-1}. \qquad (11)$$

Luego, q_j es el vector unitario $A_j/\|A_j\|$.

Observación sobre los cálculos Considero que es más fácil calcular los ortogonales a, B, C, sin forzar a que sus longitudes sean iguales a 1. Luego, las raíces cuadradas entran sólo al final, cuando se divide entre esas longitudes. El ejemplo anterior debería tener las mismas B y C, sin utilizar raíces cuadradas. Observe el $\frac{1}{2}$ de $a^{\mathrm{T}}b/a^{\mathrm{T}}a$ en vez del $\frac{1}{\sqrt{2}}$ de $q^{\mathrm{T}}b$:

$$B = \begin{bmatrix}1\\0\\0\end{bmatrix} - \frac{1}{2}\begin{bmatrix}1\\0\\1\end{bmatrix} \quad \text{y entonces} \quad C = \begin{bmatrix}2\\1\\0\end{bmatrix} - \begin{bmatrix}1\\0\\1\end{bmatrix} - 2\begin{bmatrix}\frac{1}{2}\\0\\-\frac{1}{2}\end{bmatrix}.$$

La factorización $A = QR$

Se empezó con una matriz A, cuyas columnas eran a, b, c. Se termina con una matriz Q, cuyas columnas son q_1, q_2, q_3. ¿Cuál es la relación entre estas matrices? Las matrices A y Q son de m por n cuando los n vectores están en el espacio m-dimensional, y debe haber una tercera matriz que las relacione.

La idea es escribir las as como combinaciones de las qs. El vector b en la figura 3.10 es una combinación de los q_1 y q_2 ortonormales, y se sabe de cuál combinación se trata:

$$b = (q_1^{\mathrm{T}}b)q_1 + (q_2^{\mathrm{T}}b)q_2.$$

Todo vector en el plano es la suma de sus componentes q_1 y q_2. De manera semejante, c es la suma de sus componentes q_1, q_2, q_3: $c = (q_1^{\mathrm{T}}c)q_1 + (q_2^{\mathrm{T}}c)q_2 + (q_3^{\mathrm{T}}c)q_3$. Si lo anterior se expresa en forma matricial, se tiene ***la nueva factorización A = QR***:

Factores QR $$A = \begin{bmatrix} & & \\ a & b & c \\ & & \end{bmatrix} = \begin{bmatrix} & & \\ q_1 & q_2 & q_3 \\ & & \end{bmatrix}\begin{bmatrix}q_1^{\mathrm{T}}a & q_1^{\mathrm{T}}b & q_1^{\mathrm{T}}c\\ & q_2^{\mathrm{T}}b & q_2^{\mathrm{T}}c\\ & & q_3^{\mathrm{T}}c\end{bmatrix} = QR \qquad (12)$$

¡Observe los ceros en la última matriz! R es *triangular superior* debido a la forma en que se realizó el proceso de Gram-Schmidt. Los primeros vectores a y q_1 están sobre la misma recta. Luego, q_1 y q_2 estaban en el mismo plano que a, b. Los terceros vectores, c y q_3 no se requirieron sino hasta el paso 3.

La factorización QR es como $A = LU$, excepto que el primer factor Q tiene columnas ortonormales. El segundo factor se denomina R, porque los elementos diferentes de cero están a la *derecha* de la diagonal (y la letra U ya se había utilizado). Los elementos fuera de R fuera de la diagonal son los números $q_1^{\mathrm{T}}b = 1/\sqrt{2}$ y $q_1^{\mathrm{T}}c = q_2^{\mathrm{T}}c = \sqrt{2}$, que se encontraron antes. Toda la factorización es

$$A = \begin{bmatrix}1 & 1 & 2\\ 0 & 0 & 1\\ 1 & 0 & 0\end{bmatrix} = \begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2} & 0\\ 0 & 0 & 1\\ 1/\sqrt{2} & -1/\sqrt{2} & 0\end{bmatrix}\begin{bmatrix}\sqrt{2} & 1/\sqrt{2} & \sqrt{2}\\ & 1/\sqrt{2} & \sqrt{2}\\ & & 1\end{bmatrix} = QR.$$

Las longitudes de a, B, C se observan en la diagonal de R. Los vectores ortonormales q_1, q_2, q_3, que constituyen todo el objeto de la ortogonalización, están en el primer factor Q.

Tal vez QR no es tan elegante como LU (debido a las raíces cuadradas). Ambas factorizaciones revisten una importancia crucial en la teoría del álgebra lineal, y absolutamente fundamental en los cálculos. Si LU es Hertz, entonces QR es Avis.

Los elementos $r_{ij} = q_i^{\mathrm{T}} a_j$ aparecen en la fórmula (11), cuando $\|A_j\| q_j$ se sustituye por A_j:

$$a_j = (q_1^{\mathrm{T}} a_j) q_1 + \cdots + (q_{j-1}^{\mathrm{T}} a_j) q_{j-1} + \|A_j\| q_j = Q \text{ Multiplicado por la columna } j \text{ de } R \quad (13)$$

3U Toda matriz de m por n con columnas independientes puede factorizarse en $A = QR$. Las columnas de A son ortonormales, y R es triangular superior e invertible. Cuando $m = n$, y todas las matrices son cuadradas, Q se convierte en una matriz ortogonal.

No debe olvidarse la cuestión más importante de la ortogonalización. Simplifica el problema de mínimos cuadrados a $Ax = b$. Las ecuaciones normales siguen siendo correctas, pero $A^{\mathrm{T}}A$ se vuelve más fácil:

$$A^{\mathrm{T}}A = R^{\mathrm{T}}Q^{\mathrm{T}}QR = R^{\mathrm{T}}R. \quad (14)$$

La ecuación fundamental $A^{\mathrm{T}}A\widehat{x} = A^{\mathrm{T}}b$ se simplifica a un sistema triangular:

$$R^{\mathrm{T}}R\widehat{x} = R^{\mathrm{T}}Q^{\mathrm{T}}b \quad \text{o bien} \quad R\widehat{x} = Q^{\mathrm{T}}b. \quad (15)$$

En vez de resolver $QRx = b$, lo cual no es posible, se resuelve $R\widehat{x} = Q^{\mathrm{T}}b$, lo cual supone sólo sustitución hacia atrás porque R es triangular. El costo real lo constituyen las mn^2 operaciones del proceso de Gram-Schmidt, que son necesarias para encontrar en primer lugar a Q y a R.

La misma idea de ortogonalidad se aplica a las funciones. Los senos y los cosenos son ortogonales; las potencias $1, x, x^2$ no lo son. Cuando $f(x)$ se escribe como una combinación de senos y cosenos, se trata de una ***serie de Fourier***. Cada término es una proyección sobre una recta; la recta en el espacio de funciones que contiene múltiplos de cos nx o de sen nx. Esta situación es completamente paralela al caso vectorial, y es muy importante. Por último, se tiene una tarea para Schmidt: ortogonalizar las potencias de x y producir los polinomios de Legendre.

Espacios de funciones y series de Fourier

Esta sección es breve y opcional, aunque tiene varias buenas intenciones:

1. Introducir el espacio vectorial de dimensión infinita más famoso (el *espacio de Hilbert*);
2. Extender los conceptos de longitud y producto interno de vectores v a funciones $f(x)$;
3. Identificar las series de Fourier como sumatorias de proyecciones unidimensionales (las "columnas" ortogonales son los senos y los cosenos);
4. Aplicar el proceso de ortogonalización de Gram-Schmidt a los polinomios $1, x, x^2, \ldots$; y
5. Encontrar la mejor aproximación a $f(x)$ mediante una recta.

Trataremos de seguir este esbozo, que abre una variedad de nuevas aplicaciones para el álgebra lineal de manera sistemática.

1. Espacio de Hilbert. Luego de estudiar $\mathbf{R}^n$, resulta natural pensar en el espacio $\mathbf{R}^{\infty}$. Contiene a todos los vectores $v = (v_1, v_2, v_3, \ldots)$ con una sucesión infinita de componentes. En realidad, este espacio es demasiado grande cuando no hay control sobre el tamaño de las componentes v_j. Una idea mucho mejor es preservar la conocida definición de longitud, usando una suma de cuadrados, e *incluir sólo aquellos vectores **cuya longitud sea finita***:

$$\textbf{Longitud al cuadrado} \qquad \|v\|^2 = v_1^2 + v_2^2 + v_3^2 + \cdots \quad (16)$$

La serie infinita debe converger a una suma finita. Esto deja $(1, \frac{1}{2}, \frac{1}{3}, \ldots)$ pero no $(1, 1, 1, \ldots)$. Los vectores con longitud finita pueden sumarse ($\|v + w\| \leq \|v\| + \|w\|$) y multiplicarse por escalares, de modo que forman un espacio vectorial. Es el célebre ***espacio de Hilbert***.

Este espacio es la forma natural para hacer que las dimensiones se vuelvan infinitas, y al mismo tiempo preservar la geometría del espacio euclidiano normal. Las elipses se convierten en elipsoides de dimensión infinita, y las rectas perpendiculares se reconocen exactamente como antes. Los vectores v y w son ortogonales cuando su producto interno es cero:

$$\textbf{Ortogonalidad} \qquad v^{\mathrm{T}}w = v_1w_1 + v_2w_2 + v_3w_3 + \cdots = 0.$$

Se garantiza que esta sumatoria converge, y que para dos vectores cualesquiera sigue cumpliendo la desigualdad de Schwarz $|v^{\mathrm{T}}w| \leq \|v\|\,\|w\|$. El coseno, incluso en el espacio de Hilbert, nunca es mayor que 1.

Hay otra cuestión extraordinaria respecto a este espacio: se presenta bajo numerosos disfraces. Sus "vectores" pueden ser funciones, lo cual es otra cuestión importante.

2. Longitud y productos internos. Suponga que $f(x) = \operatorname{sen} x$ sobre el intervalo $0 \leq x \leq 2\pi$. Esta f es como un vector con todo un continuo de componentes, donde los valores de sen x están a lo largo de todo el intervalo. Para encontrar la longitud de un vector así, la regla de costumbre de sumar los cuadrados de las componentes se vuelve imposible. Esta sumatoria se sustituye, de forma natural e inevitable, por la *integración*:

$$\textbf{Longitud } \|f\| \textbf{ de la función} \quad \|f\|^2 = \int_0^{2\pi} (f(x))^2\,dx = \int_0^{2\pi} (\operatorname{sen} x)^2\,dx = \pi. \quad (17)$$

El espacio de Hilbert se ha convertido en un ***espacio de funciones***. Los vectores son funciones, se cuenta con un método para medir su longitud, y el espacio contiene a todas aquéllas funciones de longitud finita, justo como en la ecuación (16). No contiene a las funciones $F(x) = 1/x$, ya que la integral de $1/x^2$ es infinita.

La misma idea de sustituir la sumatoria por la integración produce el ***producto interno de dos funciones***: Si $f(x) = \operatorname{sen} x$ y $g(x) = \cos x$, entonces su producto interno es

$$(f, g) = \int_0^{2\pi} f(x)g(x)\,dx = \int_0^{2\pi} \operatorname{sen} x \cos x\,dx = 0. \quad (18)$$

Esto es exactamente como el producto interno vectorial $f^{\mathrm{T}}g$. Sigue estando relacionado con la longitud mediante $(f, f) = \|f\|^2$. La desigualdad de Schwarz se sigue cumpliendo: $|(f, g)| \leq \|f\|\,\|g\|$. Por supuesto, dos funciones como sen x y cos x, cuyo producto interno es cero, se denominan ortogonales. Incluso, son ortogonales después de la división entre su longitud $\sqrt{\pi}$.

3. La ***serie de Fourier*** de una función es un desarrollo en senos y cosenos:

$$f(x) = a_0 + a_1 \cos x + b_1 \operatorname{sen} x + a_2 \cos 2x + b_2 \operatorname{sen} 2x + \cdots$$

Para calcular un coeficiente como b_1, ambos miembros *se multiplican* por la función correspondiente sen x y *se integra* desde 0 hasta 2π. (La función $g(x)$ está dada en ese intervalo). En otras palabras, se toma el producto interno de ambos miembros con sen x:

$$\int_0^{2\pi} f(x)\,\text{sen}\,x\,dx = a_0 \int_0^{2\pi} \text{sen}\,x\,dx + a_1 \int_0^{2\pi} \cos x\,\text{sen}\,x\,dx + b_1 \int_0^{2\pi} (\text{sen}\,x)^2\,dx + \cdots$$

En el miembro derecho, cada integral es cero excepto una: donde sen x se multiplica a sí mismo. *Los senos y los cosenos son mutuamente ortogonales*, como en la ecuación (18): Por tanto b_1 es el miembro izquierdo dividido por aquella integral no cero:

$$b_1 = \frac{\int_0^{2\pi} f(x)\,\text{sen}\,x\,dx}{\int_0^{2\pi} (\text{sen}\,x)^2\,dx} = \frac{(f, \text{sen}\,x)}{(\text{sen}\,x, \text{sen}\,x)}.$$

El coeficiente de Fourier a_1 hubiera podido ser cos x en vez de sen x, y a_2 hubiera podido ser cos $2x$.

Todo el asunto consiste en ver la analogía con las proyecciones. La componente del vector b a lo largo de la recta generada por a es $b^{\mathrm{T}}a/a^{\mathrm{T}}a$. ***Una serie de Fourier proyecta f(x) sobre sen x.*** Su componente p en esta dirección es exactamente b_1 sen x.

El coeficiente b_1 es la solución por mínimos cuadrados de la ecuación inconsistente b_1 sen $x = f(x)$. Esto lleva a b_1 sen x lo más cerca posible de $f(x)$. Todos los términos en la serie son proyecciones sobre un seno o un coseno. Debido a que los senos y los cosenos son ortogonales, *la serie de Fourier proporciona las coordenadas del "vector" f(x) respecto a un conjunto de* (una infinidad de) *ejes perpendiculares*.

4. Gram-Schmidt para funciones. Además de los senos y los cosenos, hay muchas funciones útiles que no siempre son ortogonales. Las más sencillas son las potencias de x, y lamentablemente no existe ningún intervalo sobre el cual incluso 1 y x^2 sean perpendiculares. (Su producto interno siempre es positivo, ya que es la integral de x^2.) En consecuencia, la parábola más próxima a $f(x)$ *no es* la suma de sus proyecciones sobre 1, x, y x^2. Hay una matriz como $(A^{\mathrm{T}}A)^{-1}$, y su correlato está dado por la **matriz de Hilbert** mal acondicionada. Sobre el intervalo $0 \le x \le 1$,

$$A^{\mathrm{T}}A = \begin{bmatrix} (1,1) & (1,x) & (1,x^2) \\ (x,1) & (x,x) & (x,x^2) \\ (x^2,1) & (x^2,x) & (x^2,x^2) \end{bmatrix} = \begin{bmatrix} \int 1 & \int x & \int x^2 \\ \int x & \int x^2 & \int x^3 \\ \int x^2 & \int x^3 & \int x^4 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}.$$

Esta matriz tiene una gran inversa, ya que los ejes 1, x, x^2 están lejos de ser perpendiculares. La situación se vuelve imposible si se agregan unos cuantos ejes más. *Es virtualmente inútil resolver $A^{\mathrm{T}}A\widehat{x} = A^{\mathrm{T}}b$ para el polinomio más próximo de grado diez.*

Con más precisión, resulta inútil resolver esto por eliminación gaussiana; todo error por redondeo puede amplificarse por más de 10^{13}. Por otra parte, no es posible rendirse y ya; la aproximación mediante polinomios debe ser posible. La idea correcta es cambiar a ejes ortogonales (con el proceso de Gram-Schmidt): Se buscan combinaciones de 1, x, x^2 *que sean* ortogonales.

Resulta conveniente trabajar con un intervalo situado simétricamente como $-1 \le x \le 1$, ya que así todas las potencias impares de x se hacen ortogonales a todas las potencias pares:

$$(1, x) = \int_{-1}^{1} x\,dx = 0, \qquad (x, x^2) = \int_{-1}^{1} x^3\,dx = 0.$$

En consecuencia, el proceso de Gram-Schmidt puede empezar aceptando a $v_1 = 1$ y $v_2 = x$ como los dos primeros ejes perpendiculares. Debido a que $(x, x^2) = 0$, sólo tiene los ángulos correctos entre 1 y x^2. El tercer polinomio ortogonal es

Ortogonalizar $$v_3 = x^2 - \frac{(1,x^2)}{(1,1)}1 - \frac{(x,x^2)}{(x,x)}x = x^2 - \frac{\int_{-1}^{1} x^2\,dx}{\int_{-1}^{1} 1\,dx} = x^2 - \frac{1}{3}.$$

Los polinomios que se construyen de esta manera se denominan ***polinomios de Legendre*** y son ortogonales mutuamente sobre el intervalo $-1 \le x \le 1$.

Comprobar $$\left(1, x^2 - \frac{1}{3}\right) = \int_{-1}^{1} \left(x^2 - \frac{1}{3}\right) dx = \left[\frac{x^3}{3} - \frac{x}{3}\right]_{-1}^{1} = 0.$$

Ahora ya es posible calcular el polinomio más próximo de grado 10, sin ningún desastre, al proyectar sobre los 10 (u 11) primeros polinomios de Legendre.

5. Mejor recta. Suponga que se desea aproximar $y = x^5$ mediante una recta $C + Dx$ entre $x = 0$ y $x = 1$. Hay por lo menos tres formas para encontrar esa recta, y si se comparan estas formas, entonces ¡todo el capítulo se aclara!

1. Se resuelve $\begin{bmatrix} 1 & x \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = x^5$ por mínimos cuadrados. La ecuación $A^{\mathrm{T}}A\widehat{x} = A^{\mathrm{T}}b$ es

$$\begin{bmatrix} (1,1) & (1,x) \\ (x,1) & (x,x) \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} (1,x^5) \\ (x,x^5) \end{bmatrix} \quad \text{o bien} \quad \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} \frac{1}{6} \\ \frac{1}{7} \end{bmatrix}.$$

2. Se minimiza $E^2 = \int_0^1 (x^5 - C - Dx)^2\,dx = \frac{1}{11} - \frac{2}{6}C - \frac{2}{7}D + C^2 + CD + \frac{1}{3}D^2$. Las derivadas respecto a C y D, después de dividir entre 2, traen de regreso las ecuaciones normales del método 1 (y la solución es $\widehat{C} = \frac{1}{6} - \frac{5}{14}$, $\widehat{D} = \frac{5}{7}$).

$$-\frac{1}{6} + C + \frac{1}{2}D = 0 \qquad \text{y} \qquad -\frac{1}{7} + \frac{1}{2}C + \frac{1}{3}D = 0.$$

3. Se aplica el proceso de Gram-Schmidt para sustituir x por $x - (1, x)/(1, 1)$. Esto es $x - \frac{1}{2}$, que es ortogonal a 1. Luego, las proyecciones unidimensionales producen la mejor recta:

$$C + Dx = \frac{(x^5, 1)}{(1,1)}1 + \frac{(x^5, x - \frac{1}{2})}{(x - \frac{1}{2}, x - \frac{1}{2})}\left(x - \tfrac{1}{2}\right) = \frac{1}{6} + \frac{5}{7}\left(x - \frac{1}{2}\right).$$

Conjunto de problemas 3.4

1. *a)* Escriba las cuatro ecuaciones para ajustar $y = C + Dt$ a los datos

$$y = -4 \text{ en } t = -2, \quad y = -3 \text{ en } t = -1$$
$$y = -1 \text{ en } t = 1, \quad y = 0 \text{ en } t = 2.$$

Demuestre que las columnas son ortogonales.

b) Encuentre la recta óptima, trace su gráfica, y escriba E^2.

c) Interprete el error cero en términos del sistema original de cuatro ecuaciones en dos incógnitas: el miembro derecho $(-4, -3, -1, 0)$ está en el espacio ____.

2. Proyecte $b = (0, 3, 0)$ sobre cada uno de los vectores ortogonales $a_1 = (\frac{2}{3}, \frac{2}{3}, -\frac{1}{3})$ y $a_2 = (-\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$, y luego encuentre su proyección p sobre el plano de a_1 y a_2.

3. Encuentre también la proyección de $b = (0, 3, 0)$ sobre $a_3 = \left(\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}\right)$ y sume las tres proyecciones. ¿Por qué $P = a_1a_1^T + a_2a_2^T + a_3a_3^T$ es igual a I?

4. Si Q_1 y Q_2 son matrices ortogonales, de modo que $Q^TQ = I$, demuestre que Q_1Q_2 también es ortogonal. Si Q_1 es una rotación en un ángulo θ y Q_2 es una rotación en un ángulo ϕ, ¿a qué es igual Q_1Q_2? ¿Puede encontrar las identidades trigonométricas para sen $(\theta + \phi)$ y cos $(\theta + \phi)$ en la multiplicación de matrices Q_1Q_2?

5. Si u es un vector unitario, demuestre que $Q = I - 2uu^T$ es una matriz ortogonal simétrica. (Es una reflexión, también conocida como transformación de Householder.) Calcule Q cuando $u^T = \begin{bmatrix}\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2}\end{bmatrix}$.

6. Encuentre una tercera columna de modo que la matriz

$$Q = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{14} & \\ 1/\sqrt{3} & 2/\sqrt{14} & \\ 1/\sqrt{3} & -3/\sqrt{14} & \end{bmatrix}$$

sea ortogonal. Debe ser un vector unitario que sea ortogonal a las otras columnas; ¿cuánta libertad deja esto? Compruebe que los renglones automáticamente se vuelven ortonormales al mismo tiempo.

7. Demuestre, formando directamente b^Tb, que el teorema de Pitágoras se cumple para cualquier combinación $b = x_1q_1 + \cdots + x_nq_n$ de vectores ortonormales: $\|b\|^2 = x_1^2 + \cdots + x_n^2$. En términos matriciales, $b = Qx$, de modo que esto demuestra otra vez que se preservan las longitudes: $\|Qx\|^2 = \|x\|^2$.

8. Proyecte el vector $b = (1, 2)$ sobre dos vectores que no sean ortogonales, $a_1 = (1, 0)$ y $a_2 = (1, 1)$. Demuestre que, a diferencia del caso ortogonal, la suma de las dos proyecciones unidimensionales no es igual a b.

9. Si los vectores q_1, q_2, q_3 son ortonormales, ¿cuál combinación de q_1 y q_2 es la más próxima a q_3?

10. Si q_1 y q_2 son la salida del proceso de Gram-Schmidt, ¿cuáles eran los posibles vectores de entrada a y b?

11. Demuestre que una matriz ortogonal triangular superior debe ser diagonal.

12. ¿Qué múltiplo de $a_1 = \begin{bmatrix}1\\1\end{bmatrix}$ debe restarse a $a_2 = \begin{bmatrix}4\\0\end{bmatrix}$ para que el resultado sea ortogonal a a_1? Factorice $\begin{bmatrix}1 & 4\\1 & 0\end{bmatrix}$ en QR con vectores ortonormales en Q.

13. Aplique el proceso de Gram-Schmidt a

$$a = \begin{bmatrix}0\\0\\1\end{bmatrix}, \quad b = \begin{bmatrix}0\\1\\1\end{bmatrix}, \quad c = \begin{bmatrix}1\\1\\1\end{bmatrix}$$

y escriba el resultado en la forma $A = QR$.

14. A partir de los vectores no ortogonales a, b, c, encuentre vectores ortonormales q_1, q_2, q_3:

$$a = \begin{bmatrix}1\\1\\0\end{bmatrix}, \quad b = \begin{bmatrix}1\\0\\1\end{bmatrix}, \quad c = \begin{bmatrix}0\\1\\1\end{bmatrix}.$$

15. Encuentre un conjunto ortonormal q_1, q_2, q_3 para el cual q_1 y q_2 generen el espacio columna de

$$A = \begin{bmatrix} 1 & 1\\ 2 & -1\\ -2 & 4\end{bmatrix}.$$

¿Qué subespacio fundamental contiene a q_3? ¿Cuál es la solución por mínimos cuadrados de $Ax = b$ si $b = \begin{bmatrix}1 & 2 & 7\end{bmatrix}^T$?

16. Exprese la ortogonalización de Gram-Schmidt de a_1, a_2 como $A = QR$:

$$a_1 = \begin{bmatrix}1\\2\\2\end{bmatrix}, \quad a_2 = \begin{bmatrix}1\\3\\1\end{bmatrix}.$$

Dados n vectores a_i con m componentes, ¿cuáles son las formas de A, Q, y R?

17. Con la misma matriz A del problema 16, y con $b = \begin{bmatrix}1 & 1 & 1\end{bmatrix}^T$, use $A = QR$ para resolver el problema de mínimos cuadrados $Ax = b$.

18. Si $A = QR$, encuentre una fórmula sencilla para la matriz proyección P sobre el espacio columna de A.

19. Demuestre que los siguientes *pasos modificados* del proceso de *Gram-Schmidt* producen la misma C que en la ecuación (10):

$$C^* = c - (q_1^Tc)q_1 \qquad \text{y} \qquad C = C^* - (q_2^TC^*)\,q_2.$$

Esto es mucho más estable, restar las proyecciones una a la vez.

20. En el espacio de Hilbert, encuentre la longitud del vector $v = (1/\sqrt{2}, 1/\sqrt{4}, 1/\sqrt{8}, \ldots)$ y la longitud de la función $f(x) = e^x$ (sobre el intervalo $0 \le x \le 1$). ¿Cuál es el producto interno sobre este intervalo de e^x y e^{-x}?

21. ¿Cuál es la función $a \cos x + b$ sen x más próxima a la función $f(x) =$ sen $2x$ sobre el intervalo de $-\pi$ a π? ¿Cuál es la recta $c + dx$ más próxima?

22. Iguale la derivada a cero para encontrar el valor de b_1 que minimiza

$$\|b_1 \operatorname{sen} x - \cos x\|^2 = \int_0^{2\pi} (b_1 \operatorname{sen} x - \cos x)^2\, dx.$$

Compare con el coeficiente b_1 de Fourier.

23. Encuentre los coeficientes de Fourier a_0, a_1, b_1 de la función escalón $y(x)$, que es igual a 1 sobre el intervalo $0 \le x \le \pi$ y 0 en el intervalo restante $\pi < x < 2\pi$:

$$a_0 = \frac{(y, 1)}{(1, 1)} \qquad a_1 = \frac{(y, \cos x)}{(\cos x, \cos x)} \qquad b_1 = \frac{(y, \operatorname{sen} x)}{(\operatorname{sen} x, \operatorname{sen} x)}.$$

24. Encuentre los cuatro polinomios de Legendre. Se trata de una cúbica $x^3 + ax^2 + bx + c$ que es ortogonal a 1, x, y $x^2 - \frac{1}{3}$ sobre el intervalo $-1 \le x \le 1$.

25. ¿Cuál es la recta más próxima a la parábola $y = x^2$ sobre $-1 \le x \le 1$?

26. En la fórmula de Gram-Schmidt (10), compruebe que C es ortogonal a q_1 y q_2.

27. Encuentre una base ortonormal del subespacio generado por $a_1 = (1, -1, 0, 0)$, $a_2 = (0, 1, -1, 0)$, $a_3 = (0, 0, 1, -1)$.

28. Aplique el proceso de Gram-Schmidt a $(1, -1, 0)$, $(0, 1, -1)$ y $(1, 0, -1)$ para encontrar una base ortonormal sobre el plano $x_1 + x_2 + x_3 = 0$. ¿Cuál es la dimensión de este subespacio y cuántos vectores diferentes de cero provienen del proceso de Gram-Schmidt?

29. (Recomendado) Encuentre vectores ortogonales A, B y C por Gram-Schmidt a partir de a, b, c:

$$a = (1, -1, 0, 0) \qquad b = (0, 1, -1, 0) \qquad c = (0, 0, 1, -1).$$

A, B, C y a, b, c son bases para los vectores perpendiculares a $d = (1, 1, 1, 1)$.

30. Si $A = QR$, entonces $A^TA = R^TR =$ ____ triangular multiplicada por ____ triangular. *El proceso de Gram-Schmidt sobre A corresponde a eliminación sobre* A^TA. Compare

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{con} \quad A^TA = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}.$$

Para A^TA, los pivotes son $2, \frac{3}{2}, \frac{4}{3}$ y los multiplicadores son $-\frac{1}{2}$ y $-\frac{2}{3}$.

a) Use estos multiplicadores sobre A para demostrar que la columna 1 de A y B = columna 2 $-\frac{1}{2}$ (columna 1) y C = columna 3 $-\frac{2}{3}$ (columna 2) son ortogonales.

b) Compruebe que $\|\text{columna } 1\|^2 = 2$, $\|B\|^2 = \frac{3}{2}$, y $\|C\|^2 = \frac{4}{3}$, usando los pivotes.

31. ¿Falso o verdadero? (proporcione un ejemplo en cualquier caso):
a) Q^{-1} es una matriz ortogonal cuando Q es una matriz ortogonal.
b) Si Q (de 3 por 2) tiene columnas ortonormales, entonces $\|Qx\|$ siempre es igual a $\|x\|$.

32. *a*) Encuentre una base para el subespacio $\mathbf{S}$ en $\mathbf{R}^4$ generado por todas las soluciones de

$$x_1 + x_2 + x_3 - x_4 = 0.$$

b) Encuentre una base para el complemento ortogonal $\mathbf{S}^\perp$.
c) Encuentre b_1 en $\mathbf{S}$ y b_2 en $\mathbf{S}^\perp$ de modo que $b_1 + b_2 = b = (1, 1, 1, 1)$.

3.5 LA TRANSFORMADA DISCRETA DE FOURIER

La serie de Fourier es álgebra lineal en dimensiones infinitas. Los "vectores" son funciones $f(x)$; éstas son proyectadas sobre los senos y los cosenos. Así se obtienen los coeficientes de Fourier a_k y b_k. A partir de esta serie infinita de senos y cosenos, multiplicados por a_k y b_k, es posible reconstruir a $f(x)$. Este es el caso clásico, en el que soñaba Fourier, aunque en los cálculos verdaderos lo que se calcula es la ***transformada discreta de Fourier*** (TDF). Fourier sigue vivo, pero en dimensiones finitas.

Aquí se trata de álgebra lineal pura, basada en la ortogonalidad. La entrada es una sucesión de números $y_0, \ldots, y_{n-1}$, en vez de una función $f(x)$. La salida $c_0, \ldots, c_{n-1}$ tiene la misma longitud n. La relación entre y y c es lineal, de modo que debe estar dada por una matriz. Esta es la ***matriz F de Fourier***, y toda la tecnología del procesamiento de señales digitales depende de ella. La matriz de Fourier posee propiedades extraordinarias.

Las señales se digitalizan, ya sea que provengan del habla, de imágenes, del sonar o de TV (o incluso de la exploración petrolera). Las señales son transformadas por la matriz F, y después es posible transformarlas nuevamente con la finalidad de reconstruirlas. Lo crucialmente importante es que F y F^{-1} pueden ser rápidas:

F^{-1} ***debe ser sencilla. Las multiplicaciones por F y*** F^{-1} ***deben ser rápidas.***

Estas dos afirmaciones son ciertas. F^{-1} se ha conocido durante años, y se ve justo como F. De hecho, F es simétrica y ortogonal (excepto por un factor $\sqrt{n}$), y sólo tiene un inconveniente: sus elementos son ***números complejos***. Este es un precio bajo que hay que pagar, lo cual se hace a continuación. Las dificultades son minimizadas por el hecho de que *todos los elementos de F y* F^{-1} *son potencias de un solo número* w. Este número cumple $w^n = 1$.

La transformada discreta de Fourier usa $w = i$ (y observe que $i^4 = 1$). El éxito de toda la TDF depende de la multiplicación de F por su conjugado complejo $\overline{F}$:

$$F\overline{F} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & (-i) & (-i)^2 & (-i)^3 \\ 1 & (-i)^2 & (-i)^4 & (-i)^6 \\ 1 & (-i)^3 & (-i)^6 & (-i)^9 \end{bmatrix} = 4I. \tag{1}$$

De inmediato, $F\overline{F} = 4I$ indica que $F^{-1} = \overline{F}/4$. Las columnas de F son ortogonales (para obtener los elementos cero en $4I$). Las matrices de n por n cumplen $F\overline{F} = nI$. Así, la inversa de F es justo $\overline{F}/n$. Dentro de poco se considerará el número complejo $w = e^{2\pi i/n}$ (que es igual a i para $n = 4$).

Es extraordinario que F sea tan fácil de invertir. Si esto fuese todo (y hasta 1965 *era* todo), entonces la transformada discreta tendría un lugar importante. Ahora hay más. Las multiplicaciones por F y F^{-1} pueden hacerse de una manera extremadamente rápida e ingeniosa. En vez de realizar n^2 multiplicaciones por separado, que provienen de los n^2 elementos de la matriz, para efectuar los productos matrices-vectores Fc y $F^{-1}y$, sólo se requieren $\frac{1}{2}n \log n$ pasos. Este reordenamiento de la multiplicación se denomina ***transformada rápida de Fourier*** (TRF).

Esta sección empieza con w y sus propiedades, continúa con F^{-1}, y termina con la TRF: la transformada rápida. La gran aplicación en el procesamiento de señales es la *filtración*, y la clave de su éxito es la *regla de convolución*. En lenguaje matricial, todas las "matrices circulantes" son diagonalizadas por F. Así, se reducen a dos TRFs y una matriz diagonal.

Raíces complejas de la unidad

Es posible que ecuaciones reales tengan soluciones complejas. La ecuación $x^2 + 1 = 0$ condujo a la invención de i (¡y también de $-i$!) Lo anterior fue declarado como una solución y se cerró el caso. Si alguien preguntaba sobre $x^2 - i = 0$, había una respuesta: las raíces cuadradas de un número complejo de nuevo son números complejos. Es necesario permitir las combinaciones $x + iy$, con una parte real x y una parte imaginaria y, pero ya no se requieren inventos adicionales. Todo polinomio real o complejo de grado n tiene un conjunto completo de n raíces (tal vez complejas y tal vez repetidas). Este es el teorema fundamental del álgebra.

Aquí se tiene interés en ecuaciones como $x^4 = 1$. Ésta tiene cuatro soluciones: las ***raíces cuartas de la unidad***. Las dos raíces cuadradas de la unidad son 1 y -1. Las raíces cuartas son las raíces cuadradas de las raíces cuadradas, 1 y -1, i y $-i$. El número i satisface

$i^4 = 1$ porque satisface $i^2 = -1$. Para calcular las raíces octavas de la unidad se requieren las raíces cuadradas de i, lo cual lleva a $w = (1 + i)/\sqrt{2}$. Al elevar al cuadrado w se obtiene $(1 + 2i + i^2)/2$, que es i porque $1 + i^2$ es cero. Así, $w^8 = i^4 = 1$. Aquí debe haber un sistema.

Los números complejos $\cos\theta + i \operatorname{sen}\theta$ en la matriz de Fourier son extremadamente especiales. La parte real se traza sobre el eje x y la parte imaginaria, sobre el eje y (véase la figura 3.11). Así, el número w está sobre la ***circunferencia unitaria***; su distancia al origen es $\cos^2\theta + \operatorname{sen}^2\theta = 1$. Forma un ángulo θ con la horizontal. Todo el plano se estudia en el capítulo 5, donde los números complejos aparecen como valores característicos (incluso de matrices reales). Aquí sólo se requieren puntos especiales w, todos ellos sobre la circunferencia unitaria, para resolver $w^n = 1$.

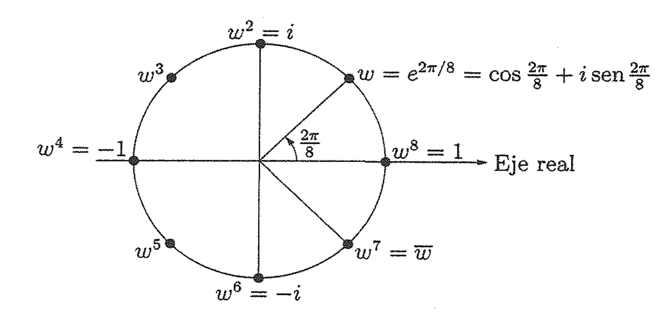


Figura 3.11 Las ocho soluciones de $z^8 = 1$ son $1, w, w^2, \ldots, w^7$ con $w = (1 + i)/\sqrt{2}$.

El cuadrado de w puede encontrarse directamente (simplemente duplica el ángulo):

$$w^2 = (\cos\theta + i \operatorname{sen}\theta)^2 = \cos^2\theta - \operatorname{sen}^2\theta + 2i \operatorname{sen}\theta\cos\theta.$$

La parte real de $\cos^2\theta - \operatorname{sen}^2\theta$ es $\cos 2\theta$, y la parte imaginaria $2 \operatorname{sen}\theta\cos\theta$ es $\operatorname{sen} 2\theta$. (Observe que no se ha incluido a i; la parte imaginaria es un número real). Por tanto, $w^2 = \cos 2\theta + i \operatorname{sen} 2\theta$. El cuadrado de w sigue estando en la circunferencia unitaria, pero ***al ángulo doble*** 2θ. Esto hace sospechar que w^n está en el ángulo $n\theta$, y la sospecha es correcta.

Hay una mejor manera de tomar potencias de w. La combinación del coseno y el seno es una exponencial compleja, con amplitud 1 y ángulo de fase θ:

$$\cos\theta + i \operatorname{sen}\theta = e^{i\theta}. \tag{2}$$

Las reglas para multiplicar, como $(e^2)(e^3) = e^5$, se siguen cumpliendo cuando los exponentes $i\theta$ son imaginarios. ***Las potencias de $w = e^{i\theta}$ permanecen en la circunferencia unitaria:***

Potencias de w
$$w^2 = e^{i2\theta}, \quad w^n = e^{in\theta}, \quad \frac{1}{w} = e^{-i\theta}. \tag{3}$$

La n-ésima potencia está al ángulo $n\theta$. Cuando $n = -1$, ***el ángulo del recíproco $1/w$ es $-\theta$***. Si $\cos\theta + i \operatorname{sen}\theta$ se multiplica por $\cos(-\theta) + i \operatorname{sen}(-\theta)$, se obtiene la respuesta 1:

$$e^{i\theta}e^{-i\theta} = (\cos\theta + i \operatorname{sen}\theta)(\cos\theta - i \operatorname{sen}\theta) = \cos^2\theta + \operatorname{sen}^2\theta = 1.$$

Nota Recuerdo el día en que al MIT llegó una carta enviada por un prisionero en Nueva York, preguntando si la fórmula de Euler (2) era cierta. Es verdaderamente sorprendente que tres de las funciones clave en matemáticas se presenten juntas de forma tan elegante. Nuestra mejor respuesta fue considerar la serie de potencias de la exponencial:

$$e^{i\theta} = 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \cdots.$$

La parte real $1 - \theta^2/2 + \cdots$ es $\cos\theta$. La parte imaginaria $\theta - \theta^3/6 + \cdots$ es el seno. La fórmula es correcta, y me hubiera gustado enviar una demostración más hermosa.

Con esta fórmula es posible resolver $w^n = 1$. Esto se convierte en $e^{in\theta} = 1$, de modo que $n\theta$ debe llevarnos alrededor de la circunferencia unitaria y volver al principio. La solución es escoger $\theta = 2\pi/n$: *la n-ésima raíz "primitiva" de la unidad es*

$$w_n = e^{2\pi i/n} = \cos\frac{2\pi}{n} + i \operatorname{sen}\frac{2\pi}{n}. \tag{4}$$

Su n-ésima potencia es $e^{2\pi i}$, que es igual a 1. Para $n = 8$, esta raíz es $(1 + i)/\sqrt{2}$:

$$w_4 = \cos\frac{\pi}{2} + i \operatorname{sen}\frac{\pi}{2} = i \quad \text{y} \quad w_8 = \cos\frac{\pi}{4} + i \operatorname{sen}\frac{\pi}{4} = \frac{1 + i}{\sqrt{2}}$$

La raíz cuarta está en $\theta = 90°$, que es $\frac{1}{4}$ (360°). Las otras raíces cuartas son las potencias $i^2 = -1$, $i^3 = -i$, e $i^4 = 1$. Las otras raíces octavas son las potencias $w_8^2, w_8^3, \ldots, w_8^8$. Las raíces son equidistantes sobre la circunferencia unitaria, a intervalos de $2\pi/n$. Observe nuevamente que el cuadrado de w_8 es w_4, lo cual es esencial en la transformada de Fourier rápida. ***La suma de las raíces es cero***. Primero $1 + i - 1 - i = 0$, y luego

Suma de las raíces octavas
$$1 + w_8 + w_8^2 + \cdots + w_8^7 = 0. \tag{5}$$

Una demostración es multiplicar el miembro izquierdo por w_8, que lo deja sin cambio. (Se obtiene $w_8 + w_8^2 + \cdots + w_8^8$, y w_8^8, es igual a 1). Cada uno de los ocho puntos se desplaza un ángulo de 45°, aunque siguen siendo los mismos ocho puntos. Debido a que cuando se multiplica por w_8 el único número que permanece sin cambio es el cero, la suma debe ser cero. Cuando n es par, las raíces se cancelan por pares (como $1 + i^2 = 0$ e $i + i^3 = 0$). Sin embargo, la suma de las tres raíces cúbicas de 1 también es igual a cero.

La matriz de Fourier y su inversa

En el caso continuo, la serie de Fourier puede reproducir a $f(x)$ sobre todo un intervalo. Utiliza una infinidad de senos y cosenos (o exponenciales). En el caso discreto, con sólo n coeficientes $c_0, \ldots, c_{n-1}$ a elegir, sólo se pide *igualdad en n puntos*. Así se obtienen n ecuaciones. Se reproducen los cuatro valores $y = 2, 4, 6, 8$ cuando $Fc = y$:

$$\boldsymbol{Fc = y} \qquad \begin{aligned} c_0 + c_1 + c_2 + c_3 &= 2 \\ c_0 + ic_1 + i^2c_2 + i^3c_3 &= 4 \\ c_0 + i^2c_1 + i^4c_2 + i^6c_3 &= 6 \\ c_0 + i^3c_1 + i^6c_2 + i^9c_3 &= 8. \end{aligned} \tag{6}$$

La secuencia de entrada es $y = 2, 4, 6, 8$. La secuencia de salida es c_0, c_1, c_2, c_3. Las cuatro ecuaciones (6) buscan una serie de Fourier de cuatro términos que corresponda a las entradas en cuatro puntos equidistantes x sobre el intervalo de 0 a 2π:

Serie de Fourier discreta

$$c_0 + c_1e^{ix} + c_2e^{2ix} + c_3e^{3ix} = \begin{matrix} 2 & \text{en} & x = 0 \\ 4 & \text{en} & x = \pi/2 \\ 6 & \text{en} & x = \pi \\ 8 & \text{en} & x = 3\pi/2. \end{matrix}$$

Éstas son las cuatro ecuaciones en el sistema (6). En $x = 2\pi$, la serie devuelve $y_0 = 2$ y continúa periódicamente. La serie de Fourier discreta se escribe mejor en su forma *compleja*, como una combinación de exponenciales e^{ikx}, en vez de como sen kx y cos kx.

Para toda n, la matriz que relaciona y con c puede invertirse. Representa n ecuaciones, lo cual requiere que la serie finita $c_0 + c_1e^{ix} + \cdots$ (***n* términos**) coincida con y (***en n puntos***). La primera coincidencia es en $x = 0$, donde $c_0 + \cdots + c_{n-1} = y_0$. Los puntos restantes conllevan potencias de w, y el problema completo es $Fc = y$:

$$Fc = y \qquad \begin{bmatrix} 1 & 1 & 1 & \cdot & 1 \\ 1 & w & w^2 & \cdot & w^{n-1} \\ 1 & w^2 & w^4 & \cdot & w^{2(n-1)} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & w^{n-1} & w^{2(n-1)} & \cdot & w^{(n-1)^2} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \cdot \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \cdot \\ y_{n-1} \end{bmatrix}. \qquad (7)$$

Ahí queda la matriz F de Fourier con elementos $F_{jk} = w^{jk}$. Resulta natural numerar los renglones y las columnas desde 0 hasta $n - 1$, en vez de hacerlo desde 1 hasta n. El primer renglón tiene $j = 0$, la primera columna tiene $k = 0$, y todos los elementos son $w^0 = 1$.

Para encontrar las cs es necesario invertir F. En el caso de 4 por 4, F^{-1} se construyó a partir de $1/i = -i$. Esta es la regla general, que F^{-1} proviene del número complejo $w^{-1} = \overline{w}$. Está en el ángulo $-2\pi/n$, donde w estaba en el ángulo $+2\pi/n$:

3V La matriz inversa se construye a partir de las potencias de $w^{-1} = 1/w = \overline{w}$:

$$F^{-1} = \frac{1}{n} \begin{bmatrix} 1 & 1 & 1 & \cdot & 1 \\ 1 & w^{-1} & w^{-2} & \cdot & w^{-(n-1)} \\ 1 & w^{-2} & w^{-4} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & w^{-(n-1)} & w^{-2(n-1)} & \cdot & w^{-(n-1)^2} \end{bmatrix} = \frac{\overline{F}}{n}. \qquad (8)$$

$$\text{Así} \quad F = \begin{bmatrix} 1 & 1 & 1 \\ 1 & e^{2\pi i/3} & e^{4\pi i/3} \\ 1 & e^{4\pi i/3} & e^{8\pi i/3} \end{bmatrix} \quad \text{tiene} \quad F^{-1} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & e^{-2\pi i/3} & e^{-4\pi i/3} \\ 1 & e^{-4\pi i/3} & e^{-8\pi i/3} \end{bmatrix}.$$

La multiplicación del renglón j de F por la columna j de F^{-1} siempre es $(1 + 1 + \cdots + 1)/n = 1$. La parte más difícil es fuera de la diagonal, demostrar que el renglón j de F multiplicado por la columna k de F^{-1} es cero:

$$1 \cdot 1 + w^j w^{-k} + w^{2j} w^{-2k} + \cdots + w^{(n-1)j} w^{-(n-1)k} = 0 \quad \text{si} \quad j \neq k. \qquad (9)$$

La clave es observar que estos términos son las potencias de $W = w^j w^{-k}$:

$$1 + W + W^2 + \cdots + W^{n-1} = 0. \qquad (10)$$

El número W sigue siendo una raíz de la unidad: $W^n = w^{nj} w^{-nk}$ es igual a $1^j 1^{-k} = 1$. Ya que j es distinto de k, W es diferente de 1. Es una de las *otras* raíces sobre la circunferencia unitaria.

Estas raíces satisfacen $1 + W + \cdots + W^{n-1} = 0$. Otra demostración proviene de

$$1 - W^n = (1 - W)(1 + W + W^2 + \cdots + W^{n-1}). \qquad (11)$$

Debido a que $W^n = 1$, el miembro izquierdo es cero. Pero W no es 1, por lo que el último factor debe ser cero. ***Las columnas de F son ortogonales.***

La transformada rápida de Fourier (TRF)

El análisis de Fourier constituye una hermosa teoría, además de ser bastante práctico. Analizar una forma de onda en sus frecuencias es la mejor manera de aislar una señal. El proceso inverso regresa la señal. Por razones físicas y matemáticas, las exponenciales son especiales, y es posible precisar una razón central: ***Si e^{ikx} se diferencia, se integra, o si x se traslada a x + h, el resultado sigue siendo un múltiplo de e^{ikx}.*** Las exponenciales se ajustan especialmente a ecuaciones diferenciales, a ecuaciones integrales, y a ecuaciones en diferencias. Cada componente de la frecuencia va a su manera, como vector característico, y se recombinan en la solución. El análisis y la síntesis de señales –calculando c a partir de y y y a partir de c– constituyen una parte toral del cálculo científico.

Se quiere demostrar que Fc y $F^{-1}y$ pueden efectuarse rápidamente. La clave está en la relación de F_4 con F_2, o mejor aún, con *dos copias* de F_2, que van en la matriz F_2^*:

$$F_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix} \quad \text{está próximo a} \quad F_2^* = \begin{bmatrix} 1 & 1 & & \\ 1 & -1 & & \\ & & 1 & 1 \\ & & 1 & -1 \end{bmatrix}.$$

F_4 contiene las potencias de $w_4 = i$, la *raíz cuarta* de 1. F_2^* contiene las potencias de $w_2 = -1$, la *raíz cuadrada* de 1. Observe especialmente que la mitad de los elementos en F_2^*. La transformada de 2 por 2, aplicada dos veces, sólo requiere la mitad de trabajo que una transformada directa de 4 por 4. Si una transformada de 64 por 64 puede sustituirse por dos transformadas de 32 por 32, entonces el trabajo se reduce a la mitad (más el costo de volver a ensamblar los resultados). Lo que hace realidad, y posible en la práctica, lo anterior, es la simple relación entre w_{64} y w_{32}:

$$(w_{64})^2 = w_{32}, \quad \text{o bien} \quad (e^{2\pi i/64})^2 = e^{2\pi i/32}.$$

La trigésimosegunda raíz está dos veces más lejos en la circunferencia unitaria que la sexagésimacuarta raíz. Si $w^{64} = 1$, entonces $(w^2)^{32} = 1$. La m-ésima raíz es el cuadrado de la n-ésima raíz, si m es la mitad de n:

$$w_n^2 = w_m \quad \text{si} \quad m = \tfrac{1}{2}n. \qquad (12)$$

La rapidez de la TRF, en la forma estándar aquí presentada, depende de trabajar con números altamente compuestos como $2^{10} = 1024$. Sin la transformada rápida, se requieren $(1024)^2$ multiplicaciones para obtener F por c (lo cual se desea a menudo). En contraste, una transformada rápida puede realizar cada multiplicación en sólo $5 \cdot 1024$ pasos. ***Es 200 veces más rápida***, ya que sustituye un factor de 1024 por 5. En general, sustituye n^2 multiplicaciones por $\frac{1}{2}n\ell$, cuando n es 2^ℓ. Al relacionar F_n con dos copias de $F_{n/2}$, y luego con cuatro copias de $F_{n/4}$, y finalmente con una F muy pequeña, los n^2 pasos de costumbre se reducen a $\frac{1}{2}\, n \log_2 n$.

Es necesario ver cómo $y = F_n c$ (un vector con n componentes) puede recuperarse a partir de dos vectores que sólo miden la mitad de largo. El primer paso es dividir c entre

sí mismo, separando sus componentes con número par de sus componentes con número impar:

$$c' = (c_0, c_2, \ldots, c_{n-2}) \qquad \text{y} \qquad c'' = (c_1, c_3, \ldots, c_{n-1}).$$

Los coeficientes simplemente van de forma alterna en c' y c''. A partir de estos vectores, la transformada a la mitad de tamaño proporciona $y' = F_m c'$ y $y'' = F_m c''$. *Estas son las dos multiplicaciones por la matriz más pequeña* F_m. El problema central consiste en recuperar y a partir de los vectores y' y y'' que miden la mitad, Cooley y Tukey se dieron cuenta de cómo hacerlo:

3W Las m primeras y las m últimas componentes del vector $y = F_n c$ son

$$\begin{aligned} y_j &= y'_j + w_n^j y''_j, \quad j = 0, \ldots, m-1 \\ y_{j+m} &= y'_j - w_n^j y''_j, \quad j = 0, \ldots, m-1. \end{aligned} \tag{13}$$

Por tanto, los tres pasos son: separar c en c' y c''; transformarlos por medio de F_m en y' y y''; y reconstruir y a partir de la ecuación (13).

En un momento se comprobará que haciendo lo anterior se obtiene la y correcta. (Quizá el lector prefiera un diagrama de flujo al álgebra). ***Esta idea puede repetirse. Se va de*** F_{1024} ***a*** F_{512} ***a*** F_{256}. La cuenta final es $\frac{1}{2}n\ell$, cuando se empieza con la potencia $n = 2^\ell$ y todo el camino la dirección es hacia $n = 1$, donde no se requiere multiplicación. Este número $\frac{1}{2}n\ell$, satisface la regla antes proporcionada: *el doble del conteo para m, más m multiplicaciones extra, producen el conteo para n*:

$$2\left(\frac{1}{2}m(\ell-1)\right) + m = \frac{1}{2}n\ell.$$

Otra forma de contar: De $n = 2^\ell$ a $n = 1$ hay ℓ pasos. Cada paso requiere $n/2$ multiplicaciones por $D_{n/2}$ en la ecuación (13), que en realidad es una factorización de F_n:

Un paso de la TRF

$$F_{1024} = \begin{bmatrix} I_{512} & D_{512} \\ I_{512} & -D_{512} \end{bmatrix} \begin{bmatrix} F_{512} & \\ & F_{512} \end{bmatrix} \begin{bmatrix} \text{permutación} \\ \text{par-impar} \end{bmatrix}. \tag{14}$$

El costo es sólo ligeramente más que lineal. El análisis de Fourier ha sido transformado completamente por la TFR. Para comprobar la ecuación (13), es necesario separar y_j en *par* e *impar*:

$$y_j = \sum_{k=0}^{n-1} w_n^{jk} c_k \quad \text{es idéntica a} \quad \sum_{k=0}^{m-1} w_n^{2kj} c_{2k} + \sum_{k=0}^{m-1} w_n^{(2k+1)j} c_{2k+1}.$$

Cada sumatoria de la derecha consta de $m = \frac{1}{2}n$ términos. Debido a que w_n^2 es w_m, las dos sumatorias son:

$$y_j = \sum_{k=0}^{m-1} w_m^{kj} c'_k + w_n^j \sum_{k=0}^{m-1} w_m^{kj} c''_k = y'_j + w_n^j y''_j. \tag{15}$$

Para la segunda parte de la ecuación (13), $j + m$ en vez de j produce un cambio de signo:

Dentro de las sumatorias, $w_m^{k(j+m)}$ permanece w_m^{kj} debido a que $w_m^{km} = 1^k = 1$.

Fuera $w_n^{j+m} = -w_n^j$ ya que $w_n^m = e^{2\pi i m/n} = e^{\pi i} = -1$.

La idea de la TRF se modifica fácilmente con la finalidad de permitir otros factores primos de n (no sólo potencias de 2). Si n mismo es primo, se aplica un algoritmo completamente distinto.

Ejemplo 1 Los pasos de $n = 4$ a $m = 2$ son

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} \to \begin{bmatrix} c_0 \\ c_2 \\ c_1 \\ c_3 \end{bmatrix} \to \begin{bmatrix} F_2 c' \\ \\ F_2 c'' \end{bmatrix} \to \begin{bmatrix} \\ y \\ \\ \end{bmatrix}.$$

Combinados, los tres pasos multiplican c por F_4 para obtener y. Debido a que cada paso es lineal, debe provenir de una matriz, y el producto de estas matrices debe ser F_4:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix} = \begin{bmatrix} 1 & & 1 & \\ & 1 & & i \\ 1 & & -1 & \\ & 1 & & -i \end{bmatrix} \begin{bmatrix} 1 & 1 & & \\ 1 & -1 & & \\ & & 1 & 1 \\ & & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & & 1 & \\ & 1 & & \\ & & & 1 \end{bmatrix}. \tag{16}$$

Las dos copias de F_2 se reconocen en el centro. A la derecha se encuentra la matriz permutación que separa c en c' y c''. A la izquierda está la matriz que multiplica por w_n^j. Si se hubiera empezado con F_8, la matriz de en medio hubiera contenido dos copias de F_4. ***Cada una de éstas se hubiera separado como se hizo arriba.*** Entonces, ¡la TRF representa una factorización gigante de la matriz de Fourier! La simple matriz con n^2 no ceros es un producto de aproximadamente $\ell = \log_2 n$ matrices (y una permutación) con un total de sólo $n\ell$ no ceros.

La TRF y la mariposa

El primer paso de la TRF cambia la multiplicación por F_n en dos multiplicaciones por F_m. Las componentes con número par (c_0, c_2) se transforman por separado de (c_1, c_3). En la figura 3.12 se proporciona un diagrama de flujo para $n = 4$.

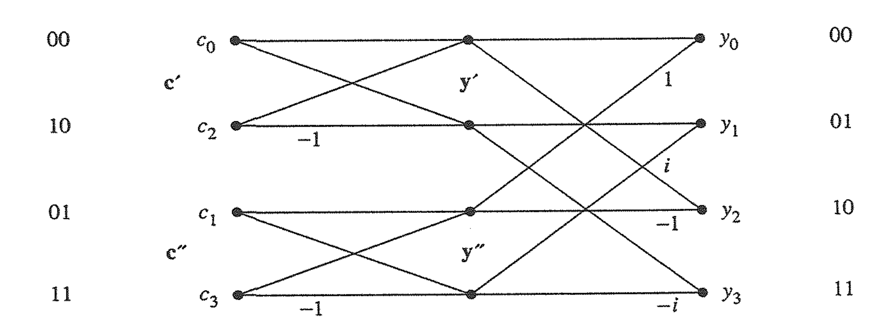


Figura 3.12 Diagrama de flujo para la transformada de Fourier rápida con $n = 4$.

Para $n = 8$, la idea clave es ***sustituir cada recuadro*** F_4 ***por recuadros*** F_2. El nuevo factor $w_4 = i$ es el cuadrado del factor anterior $w = w_8 = e^{2\pi i/8}$. El diagrama de flujo muestra el orden en que las cs entran en la TRF y las $\log_2 n$ etapas que les lleva entrar, y también muestra la sencillez de la lógica.

Cada etapa requiere $\frac{1}{2}n$ multiplicaciones, de modo que el conteo final es $\frac{1}{2}n \log n$. Hay una regla asombrosa para la permutación global de cs antes de entrar en la TRF: Escribir los subíndices $0, \ldots, 7$ en binario e *invertir el orden de sus bits*. Los subíndices aparecen en "orden invertido de bits" en la parte izquierda de la gráfica. Los números pares aparecen antes que los impares (los números que terminan con 0 aparecen primero que los números que terminan con 1).

Conjunto de problemas 3.5

1. ¿Qué son F^2 y F^4 para la matriz F de Fourier de 4 por 4?

2. Encuentre una permutación P de las columnas de F que produzca $FP = \overline{F}$ (de n por n). Combine lo anterior con $F\overline{F} = nI$ con la finalidad de encontrar F^2 y F^4 para la matriz F de Fourier de n por n.

3. Si se forma una submatriz de 3 por 3 de la matriz F_6 de 6 por 6, manteniendo sólo los elementos que están en sus renglones y columnas, primero(a), tercero(a) y cuarto(a), ¿cuál es esa submatriz?

4. Indique todas las raíces sextas de 1 en el plano complejo. ¿Cuál es la raíz primitiva w_6? (Encuentre sus partes real e imaginaria). ¿Qué potencia de w_6 es igual a $1/w_6$? ¿Cuál es $1 + w + w^2 + w^3 + w^4 + w^5$?

5. Encuentre todas las soluciones de la ecuación $e^{ix} = -1$, y todas las soluciones de $e^{i\theta} = i$.

6. ¿Cuáles son el cuadrado y la raíz cuadrada de w_{128}, la centésimovigesimaoctava raíz primitiva de 1?

7. Resuelva el sistema (6) de 4 por 4 si los miembros derechos son $y_0 = 2, y_1 = 0, y_2 = 2, y_3 = 0$. En otras palabras, resuelva $F_4c = y$.

8. Resuelva el mismo sistema con $y = (2, 0, -2, 0)$ sabiendo que F_4^{-1} y calculando $c = F_4^{-1}$. Compruebe que $c_0 + c_1e^{ix} + c_2e^{2ix} + c_3e^{3ix}$ asume los valores $2, 0, -2, 0$ en los puntos $x = 0, \pi/2, \pi, 3\pi/2$.

9. *a)* Si $y = (1, 1, 1, 1)$, demuestre que $c = (1, 0, 0, 0)$ satisface $F_4c = y$.
   b) Ahora suponga que $y = (1, 0, 0, 0)$, y encuentre c.

10. Para $n = 2$, escriba y_0 a partir de la primera línea de la ecuación (13) y y_1 a partir de la segunda línea. Para $n = 4$, use la primera línea para encontrar y_0 y y_1 y la segunda línea para encontrar y_2 y y_3, todo en términos de y' y y''.

11. Calcule $y = F_4c$ mediante los tres pasos de la transformada rápida de Fourier si $c = (1, 0, 1, 0)$. Repita el cálculo con $c = (0,1,0,1,0,1,0,1)$.

12. Calcule $y = F_8c$ mediante los tres pasos de la transformada rápida de Fourier si $c = (1, 0, 1, 0, 1, 0, 1, 0)$. Repita el cálculo con $c = (0, 1, 0, 1, 0, 1, 0, 1)$.

13. Para la matriz de 4 por 4, escriba las fórmulas para c_0, c_1, c_2, c_3 y compruebe que *si f es impar, entonces c es impar*. El vector f es impar si $f_{n-j} = -f_j$; para $n = 4$ esto significa $f_0 = 0, f_3 = -f_1, f_2 = 0$ como en sen 0, sen $\pi/2$, sen π, sen $3\pi/2$. Esto es copiado por c y conduce a una transformación rápida del seno.

14. Multiplique las tres matrices en la ecuación (16) y compare con F. ¿En cuáles seis elementos es necesario saber que $i^2 = -1$?

15. Invierta los tres factores en la ecuación (14) para encontrar una factorización rápida de F^{-1}.

16. F es simétrica. Así, ¡trasponga la ecuación (14) para encontrar una nueva transformada rápida de Fourier!

17. Todos los elementos en la factorización de F_6 implican potencias de w = raíz sexta de 1:

$$F_6 = \begin{bmatrix} I & D \\ I & -D \end{bmatrix} \begin{bmatrix} F_3 & \\ & F_3 \end{bmatrix} \begin{bmatrix} P \end{bmatrix}.$$

Escriba estos factores con $1, w, w^2$ en D y $1, w^2, w^4$ en F_3. ¡Multiplique!

En los problemas 18 a 20 se introducen los conceptos de vector característico y valor característico, cuando el producto de una matriz multiplicada por un vector es un múltiplo de ese vector. Éste es el tema del capítulo 5.

18. Las columnas de la matriz de Fourier F son los *valores característicos* de la permutación cíclica P. Multiplique PF para encontrar los valores característicos λ_0 a λ_3:

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix} \begin{bmatrix} \lambda_0 & & & \\ & \lambda_1 & & \\ & & \lambda_2 & \\ & & & \lambda_3 \end{bmatrix}.$$

Ésta es $PF = F\Lambda$, o $P = F\Lambda F^{-1}$.

19. Dos vectores característicos de esta matriz circulante c son $(1, 1, 1, 1)$ y $(1, i, i^2, i^3)$. ¿Cuáles son los valores característicos e_0 y e_1?

$$\begin{bmatrix} c_0 & c_1 & c_2 & c_3 \\ c_3 & c_0 & c_1 & c_2 \\ c_2 & c_3 & c_0 & c_1 \\ c_1 & c_2 & c_3 & c_0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = e_0 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{y} \quad C \begin{bmatrix} 1 \\ i \\ i^2 \\ i^3 \end{bmatrix} = e_1 \begin{bmatrix} 1 \\ i \\ i^2 \\ i^3 \end{bmatrix}.$$

20. Encuentre los valores característicos de la matriz "periódica" C −1, 2, −1. Los 1s en las esquinas de C la hacen periódica (**una matriz circulante**):

$$C = \begin{bmatrix} 2 & -1 & 0 & -1 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ -1 & 0 & -1 & 2 \end{bmatrix} \quad \text{tiene } c_0 = 2,\ c_1 = -1,\ c_2 = 0,\ c_3 = -1.$$

21. Para multiplicar C por x, cuando $C = FEF^{-1}$, es posible multiplicar $F(E(F^{-1}x))$. En el producto directo Cx se requieren n^2 multiplicaciones por separado. Si se conocen E y F, con el segundo método sólo se usan $n \log_2 n + n$ multiplicaciones. ¿Cuántas de éstas provienen de F, y cuántas de F^{-1}?

22. ¿Cómo podría calcular rápidamente estas cuatro componentes de Fc empezando con $c_0 + c_2, c_0 - c_2, c_1 + c_3, c_1 - c_3$? ¡Está encontrando la transformada rápida de Fourier!

$$Fc = \begin{bmatrix} c_0 + c_1 + c_2 + c_3 \\ c_0 + ic_1 + i^2c_2 + i^3c_3 \\ c_0 + i^2c_1 + i^4c_2 + i^6c_3 \\ c_0 + i^3c_1 + i^6c_2 + i^9c_3 \end{bmatrix}.$$

Capítulo 3 Ejercicios de repaso

3.1 Encuentre la longitud de $a = (2, -2, 1)$, y escriba dos vectores independientes que sean perpendiculares a a.

3.2 Encuentre todos los vectores que sean perpendiculares a $(1, 3, 1)$ y $(2, 7, 2)$, haciendo a éstos los renglones de A y resolviendo $Ax = 0$.

3.3 ¿Cuál es el ángulo entre $a = (2, -2, 1)$ y $b = (1, 2, 2)$?

3.4 ¿Cuál es la proyección p de $b = (1, 2, 2)$ sobre $a = (2, -2, 1)$?

3.5 Encuentre el coseno del ángulo entre los vectores $(3, 4)$ y $(4, 3)$.

3.6 ¿Dónde está la proyección de $b = (1, 1, 1)$ sobre el plano generado por $(1, 0, 0)$ y $(1, 1, 0)$?

3.7 El sistema $Ax = b$ tiene una solución si y sólo si b es ortogonal a ¿cuál de los cuatro subespacios fundamentales?

3.8 ¿Qué recta proporciona el mejor ajuste a los datos siguientes: $b = 0$ en $t = 0$, $b = 0$ en $t = 1$, $b = 12$ en $t = 3$?

3.9 Construya la matriz proyección P sobre el espacio generado por $(1, 1, 1)$ y $(0, 1, 3)$.

3.10 ¿Cuál función constante está más próxima de $y = x^4$ (en el sentido de mínimos cuadrados) sobre el intervalo $0 \leq x \leq 1$?

3.11 Si Q es ortogonal, ¿es cierto lo mismo para Q^3?

3.12 Encuentre todas las matrices ortogonales de 3 por 3 cuyos elementos sean ceros y unos.

3.13 ¿Qué múltiplo de a_1 debe restarse a a_2, para que el resultado sea ortogonal a a_1? Dibuje una figura.

3.14 Factorice

$$\begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & 0 \end{bmatrix}$$

como QR, reconociendo que la primera columna ya es un vector unitario.

3.15 Si todo elemento de una matriz ortogonal es $\frac{1}{4}$ o $-\frac{1}{4}$, ¿cuál es el tamaño de esta matriz?

3.16 Suponga que los vectores $q_1, \ldots, q_n$ son ortonormales. Si $b = c_1q_1 + \cdots + c_nq_n$, proporcione una fórmula para calcular el primer coeficiente c_1 en términos de b y las qs.

3.17 ¿Qué palabras describen a la ecuación $A^TA\widehat{x} = A^Tb$, al vector $p = A\widehat{x} = Pb$, y a la matriz $P = A(A^TA)^{-1}A^T$?

3.18 Si los vectores ortonormales $q_1 = \left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}\right)$ y $q_2 = \left(-\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)$ son las columnas de Q, ¿cuáles son las matrices Q^TQ y QQ^T? Demuestre que QQ^T es una matriz proyección (sobre el plano de q_1 y q_2).

3.19 Si $v_1, \ldots, v_n$ es una base ortonormal de $\mathbf{R}^n$, demuestre que $v_1v_1^T + \cdots + v_nv_n^T = I$.

3.20 *¿Falso o verdadero?* Si los vectores x y y son ortogonales, y P es una proyección, entonces Px y Py son ortogonales.

3.21 Intente ajustar una recta $b = C + Dt$ que pasa por los puntos $b = 0$, $t = 2$, y $b = 6$, $t = 2$, y demuestre que las ecuaciones normales fracasan. Dibuje todas las rectas óptimas, minimizando la suma de los cuadrados de los dos errores.

3.22 ¿Cuál punto en el plano $x + y - z = 0$ es el más próximo a $b = (2, 1, 0)$?

3.23 Encuentre una base ortonormal de $\mathbf{R}^3$, empezando con el vector $(1, 1, -1)$.

3.24 Rastreadores CT examinan a un paciente desde direcciones distintas y producen una matriz que proporciona las densidades del hueso y el tejido en cada punto. Matemáticamente, el problema consiste en recuperar una matriz a partir de sus proyecciones. En el caso de 2 por 2, ¿es posible recuperar la matriz A si se conoce la suma a lo largo de cada renglón y de cada columna?

3.25 ¿Es posible recuperar una matriz de 3 por 3 si se conocen las sumas en los renglones, las sumas en las columnas y también la suma a lo largo de la diagonal principal, así como a lo largo de las otras cuatro diagonales paralelas?

3.26 Encuentre una base ortonormal del plano $x - y + z = 0$, y encuentre la matriz P que proyecte sobre el plano. ¿Cuál es el espacio nulo de P?

3.27 Sea $A = [3\ 1\ -1]$, y sea $\mathbf{V}$ el espacio nulo de A.
a) Encuentre una base para $\mathbf{V}$ y una base para $\mathbf{V}^{\perp}$.
b) Escriba una base ortonormal para $\mathbf{V}^{\perp}$, y encuentre la matriz proyección P_1 que proyecta vectores de $\mathbf{R}^3$ sobre $\mathbf{V}^{\perp}$.
c) Encuentre la matriz proyección P_2 que proyecta vectores de $\mathbf{R}^3$ sobre $\mathbf{V}$.

3.28 Aplique el proceso de Gram-Schmidt para construir un par ortonormal q_1, q_2 a partir de $a_1 = (4, 5, 2, 2)$ y $a_2 = (1, 2, 0, 0)$. Exprese a_1 y a_2 como una combinación de q_1 y q_2, y encuentre la matriz triangular R en la forma $A = QR$.

3.29 Para A, b, x, y y cualesquiera, demuestre que
a) Si $Ax = b$ y $y^TA = 0$, entonces $y^Tb = 0$.
b) Si $Ax = 0$ y $A^Ty = b$, entonces $x^Tb = 0$.
¿Cuál teorema demuestra esto sobre los subespacios fundamentales?

3.30 ¿Existe alguna matriz cuyo espacio renglón contenga a $(1, 1, 0)$ y cuyo espacio nulo contenga a $(0, 1, 1)$?

3.31 La distancia de un plano $a^Tx = c$ (en el espacio m-dimensional) al origen es $|c|/\|a\|$. ¿Cuán lejos está el plano del origen $x_1 + x_2 - x_3 - x_4 = 8$, y cuál punto en éste es el más próximo?

3.32 En el paralelogramo con vértices en 0, v, w, y $v + w$, demuestre que la suma de las longitudes al cuadrado de los cuatro lados es igual a la suma de las longitudes al cuadrado de las dos diagonales.

3.33 *a*) Encuentre una base ortonormal para el espacio columna de A.

$$A = \begin{bmatrix} 1 & -6 \\ 3 & 6 \\ 4 & 8 \\ 5 & 0 \\ 7 & 8 \end{bmatrix}.$$

b) Escriba A como QR, donde Q tiene columnas ortonormales y R es triangular superior.

c) Encuentre la solución por mínimos cuadrados de $Ax = b$, si $b = (-3, 7, 1, 0, 4)$.

3.34 Con la matriz de ponderación $W = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$, ¿cuál es al producto interno de $(1, 0)$ con $(0, 1)$?

3.35 Para resolver un sistema rectangular $Ax = b$, reemplazamos A^{-1} (que no existe) se sustituye por $(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ (que existe si las columnas de A son independientes). Demuestre que ésta es una inversa izquierda de A pero no una inversa derecha. A la izquierda de A proporciona la identidad; a la derecha proporciona la proyección P.

3.36 Encuentre la recta $C + Dt$ que ajuste mejor las mediciones $b = 0, 1, 2, 5$ en los instantes $t = 0, 1, 3, 4$.

3.37 Encuentre la curva $y = C + D2^t$ que proporcione el mejor ajuste por mínimos cuadrados a las mediciones $y = 6$ en $t = 0$, $y = 4$ en $t = 1$, y $y = 0$ en $t = 2$. Escriba las tres ecuaciones que se resuelven si la curva pasa por los tres puntos, y encuentre los mejores C y D.

3.38 Si las columnas de A son ortogonales entre sí, ¿qué puede decir sobre la forma de $A^{\mathrm{T}}A$? Si las columnas son ortonormales, ¿qué puede decir entonces?

3.39 ¿Qué condición deben cumplir las columnas de A (que puede ser rectangular) para que $A^{\mathrm{T}}A$ sea invertible?

Capítulo



4 Determinantes

4.1 INTRODUCCIÓN

Los determinantes están mucho más lejos del centro del álgebra lineal de lo que estaban hace cien años. ¡Las matemáticas siguen cambiando de dirección! Después de todo, un *simple número puede decir tanto sobre una matriz*. De todos modos, es extraordinario cuánto puede hacer este número.

Un punto de vista es: el determinante constituye una fórmula "explícita" para cada elemento de A^{-1} y $A^{-1}b$. Esta fórmula no modifica la manera de realizar los cálculos; incluso el determinante en sí se encuentra por eliminación. De hecho, la eliminación puede considerarse como la manera más eficiente de sustituir los elementos de una matriz de n por n en la fórmula. Lo que hace la fórmula es mostrar cómo A^{-1} depende de los n^2 elementos de A, y la forma en que varía cuando los elementos cambian.

Los usos más importantes de los determinantes pueden enumerarse como se muestra a continuación:

1. Prueban la invertibilidad. ***Si el determinante de A es cero, entones A es singular. Si det A ≠ 0, entonces A es invertible*** (y A^{-1} implica $1/\det A$).

La aplicación más importante, y la razón por la que este capítulo es esencial para el libro, es en la familia de matrices $A - \lambda I$. El parámetro λ se resta a lo largo de toda la diagonal principal, y el problema consiste en encontrar los *valores característicos* para los cuales $A - \lambda I$ es singular. La prueba es $\det(A - \lambda I) = 0$. Este polinomio de grado n en λ tiene exactamente n raíces. La matriz tiene n valores característicos. Este hecho se concluye por la fórmula del determinante, y no por computadora.

2. El determinante de A es igual al ***volumen*** de una caja en el espacio n-dimensional. Las aristas de la caja provienen de los renglones de A (véase la figura 4.1). Con las columnas de A se obtiene una caja totalmente diferente con el mismo volumen.

La caja más sencilla es un cubo pequeño $dV = dx\,dy\,dz$, como en $\iiint f(x, y, z)\,dV$. Suponga que se cambia a coordenadas cilíndricas mediante $x = r\cos\theta$, $y = r\operatorname{sen}\theta$, y $z = z$. También que el pequeño intervalo dx se alarga a $(dx/du)du$ —cuando u sustituye a x en una simple integral— así el elemento de volumen se convierte en $J\,dr\,d\theta\,dz$. El *determinante jacobiano* es el análogo tridimensional del factor de alargamiento dx/du:

$$\textbf{Jacobiano} \qquad J = \begin{vmatrix} \partial x/\partial r & \partial x/\partial\theta & \partial x/\partial z \\ \partial y/\partial r & \partial y/\partial\theta & \partial y/\partial z \\ \partial z/\partial r & \partial z/\partial\theta & \partial z/\partial z \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix}.$$

El valor de este determinante es $J = r$. Es la r en el elemento de volumen cilíndrico $r\,dr\,d\theta\,dz$; este elemento es nuestra pequeña caja. (Al intentar dibujarla se ve curva, aunque quizá se haga más recta a medida que las aristas se vuelven infinitesimales.)

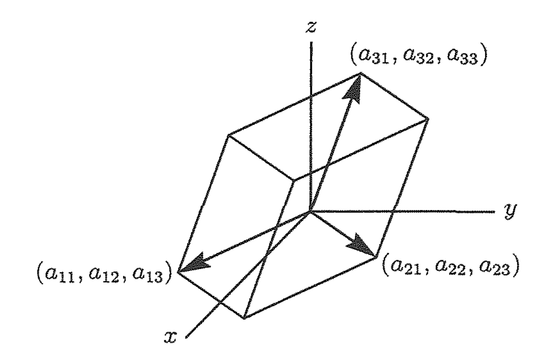


Figura 4.1 Caja formada a partir de los renglones de A: volumen = |determinante|.

3. El determinante proporciona una fórmula para cada pivote. Teóricamente, es posible pronosticar cuándo un elemento pivote es cero, lo cual requiere un intercambio de renglones. A partir de la fórmula ***determinante*** = ± (***producto de los pivotes***), se concluye que *sin importar el orden de la eliminación, el producto de los pivotes permanece igual, salvo el signo*.

Hace años, este hecho condujo a la creencia de que era inútil escapar a un pivote pequeño intercambiando renglones, ya que al final el pivote pequeño ganaría la partida. Sin embargo, lo que suele ocurrir en la práctica, en caso de que no sea posible evitar un pivote anormalmente pequeño, es que en muy poco tiempo se presenta un pivote anormalmente grande. Así es como el producto se vuelve nuevamente normal, aunque deja en ruinas la solución numérica.

4. El determinante mide la dependencia de $A^{-1}b$ respecto a cada elemento de b. Si en un experimento se cambia un parámetro, o una observación es corregida, el "coeficiente contaminante" en A^{-1} es un cociente de determinantes.

Hay otro problema respecto al determinante. Es difícil no sólo decidir acerca de su importancia, así como de su lugar idóneo en el álgebra lineal, sino también elegir la mejor definición. Resulta evidente que, det A no es alguna función extremadamente sencilla de n^2 variables; en caso contrario, sería mucho más fácil encontrar A^{-1} de lo que realmente es.

Las cuestiones sencillas sobre el determinante no son las fórmulas explícitas, sino las propiedades que posee. Esto sugiere un lugar natural para comenzar. El determinante puede (como será el caso) definirse mediante sus tres propiedades más importantes: det $I = 1$, **el signo se invierte por un intercambio de renglones, el determinante es lineal en cada renglón por separado.** Entonces el problema es demostrar cómo, usando estas propiedades de manera sistemática, es posible calcular el determinante. Esta situación retrotrae al producto de los pivotes.

En la sección 4.2 se explicarán estas tres propiedades definitorias del determinante, así como sus consecuencias más importantes. En la sección 4.3 se proporcionarán dos fórmulas más para calcular el determinante: la "gran fórmula" con $n!$ términos y una fórmula por "inducción". En la sección 4.4, el determinante se aplica para encontrar A^{-1}. Luego, $x = A^{-1}b$ se calcula con la ***regla de Cramer***. Y finalmente, en una observación opcional sobre permutaciones, se demuestra que sin importar el orden en que se apliquen las propiedades, el resultado siempre es el mismo; es decir, las propiedades definitorias son autoconsistentes.

Esta es una pregunta optimista sobre las permutaciones: ***¿cuántos cambios son necesarios para cambiar VISA en AVIS?*** Esta permutación, ¿es par o impar?

4.2 PROPIEDADES DEL DETERMINANTE

Consisten en una lista bastante larga. Por fortuna, cada regla es fácil de encontrar, e incluso más fácil de ilustrar, mediante un ejemplo de 2 por 2. En consecuencia, se comprobará que la conocida definición en el caso de 2 por 2,

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc,$$

posee cada propiedad de la lista. (Observe las dos notaciones aceptadas para el determinante, det A y $|A|$.) Las propiedades de 4 a 10 se deducen de las propiedades previas. **Cada propiedad es una consecuencia de las tres primeras.** Se recalca que las reglas son válidas para matrices cuadradas de *cualquier tamaño*.

1. *El determinante de la matriz identidad es* 1.

$$\mathbf{det}\, I = 1 \qquad \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \quad \text{y} \quad \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \quad \text{y} \ldots$$

2. *El determinante cambia de signo cuando se intercambian dos renglones.*

$$\textbf{Intercambio de renglones} \qquad \begin{vmatrix} c & d \\ a & b \end{vmatrix} = cb - ad = -\begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

El determinante de cualquier *matriz permutación* es det $P = \pm 1$. Mediante intercambio de renglones es posible transformar P en la matriz identidad. Cada intercambio de renglones cambia el signo del determinante, hasta que se llega a det $I = 1$. ¡De ahí se obtienen todas las demás matrices!

3. *El determinante depende linealmente del primer renglón.* Suponga que A, B, C son iguales a partir del segundo renglón, y que el renglón 1 de A es una combinación lineal de los primeros renglones de B y C. Entonces la regla establece: det A *es la misma combinación que* det B *y* det C.

Las combinaciones lineales implican dos operaciones: sumar vectores y multiplicar por escalares. Por consiguiente, esta regla puede separarse en dos partes:

$$\textbf{Sumar los vectores en el renglón 1} \qquad \begin{vmatrix} a+a' & b+b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}.$$

$$\textbf{Multiplicar por } t \textbf{ en el renglón 1} \qquad \begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t\begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Observe que la primera parte *no* es la afirmación falsa $\det(B + C) = \det B + \det C$. No es posible sumar todos los renglones: sólo se permite cambiar a un renglón. Ambos miembros proporcionan la respuesta $ad + a'd - bc - b'c$.

La segunda parte no es la afirmación falsa $\det(tA) = t \det A$. La matriz A tiene un factor t en *cada* renglón (y el determinante se multiplica por t^n). Es como el volumen de una caja, cuando todos los lados se alargan por 4. En n dimensiones, el volumen y el determinante se alargan por 4^n. Si sólo se alarga un lado, el volumen y el determinante se alargan por 4; ésta es la regla 3. Por la regla 2, no hay nada especial respecto al renglón 1.

Ahora ya se ha establecido el determinante, aunque este hecho no es nada evidente. En consecuencia, las reglas se aplican gradualmente para encontrar el determinante de cualquier matriz.

4. *Si dos renglones de A son iguales, entonces* $\det A = 0$.

$$\textbf{Renglones iguales} \qquad \begin{vmatrix} a & b \\ a & b \end{vmatrix} = ab - ba = 0.$$

Esto se concluye de la regla 2, ya que si se intercambian los renglones iguales, se supone que el determinante cambia de signo. Pero también queda igual, porque la matriz no cambia. El único número capaz de hacer esto es el cero, de modo que $\det A = 0$. (El razonamiento falla si $1 = -1$, que es el caso en álgebra booleana. Así, la regla 4 sustituye a la regla 2 como una de las propiedades definitorias).

5. *Restar un múltiplo de un renglón de otro renglón deja igual al determinante.*

$$\textbf{Operación en los renglones} \qquad \begin{vmatrix} a - \ell c & b - \ell d \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

La regla 3 indica que hay otro término $-\ell \begin{vmatrix} c & d \\ c & d \end{vmatrix}$, pero por la regla 4 este término es cero. ¡El paso de eliminación de costumbre no afecta al determinante!

6. *Si A tiene un renglón de ceros, entonces* $\det A = 0$.

$$\textbf{Renglón cero} \qquad \begin{vmatrix} 0 & 0 \\ c & d \end{vmatrix} = 0.$$

Una demostración consiste en sumar otro renglón al renglón cero. Por la regla 5, el determinante permanece sin cambio. Debido a que ahora la matriz cuenta con dos renglones idénticos, por la regla 4 se tiene que $\det A = 0$.

7. *Si A es triangular, entonces* $\det A$ *es el producto* $a_{11}a_{22}\cdots a_{nn}$ *de los elementos en la diagonal. Si la A triangular tiene 1s a lo largo de la diagonal, entonces* $\det A = 1$.

$$\textbf{Matriz triangular} \qquad \begin{vmatrix} a & b \\ 0 & d \end{vmatrix} = ad \qquad \begin{vmatrix} a & 0 \\ c & d \end{vmatrix} = ad.$$

Demostración Suponga que los elementos en la diagonal son diferentes de cero. Entonces la eliminación es capaz de eliminar todos los elementos fuera de la diagonal sin modificar el determinante (por la regla 5). Si A es triangular inferior, los pasos son hacia abajo, como de costumbre. Si A es triangular superior, primero se trabaja con la *última* columna, utilizando múltiplos de a_{nn}. De cualquier forma se llega a la matriz diagonal D:

$$D = \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{bmatrix} \quad \text{tiene} \quad \det D = a_{11}a_{22}\cdots a_{nn} \det I = a_{11}a_{22}\cdots a_{nn}.$$

Para encontrar $\det D$, paciente se aplica la regla 3. Al factorizar a_{11} y luego a_{22} y por último a_{nn}, se obtiene la matriz identidad. Por fin tenemos una aplicación para la regla 1: $\det I = 1$. ■

Si un elemento en la diagonal es cero, entonces la eliminación produce un renglón cero. Por la regla 5, estos pasos de la eliminación no cambian el determinante. Por la regla 6, el renglón cero significa un determinante cero. Es decir: cuando una matriz triangular es *singular* (debido a un cero en la diagonal principal), su determinante es *cero*.

Esta propiedad es fundamental. **El determinante de todas las matrices singulares es cero.**

8. *Si A es singular, entonces* $\det A = 0$. *Si A es invertible, entonces* $\det A \neq 0$.

$$\textbf{Matriz singular} \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{no es invertible si y sólo si} \quad ad - bc = 0.$$

Si A es singular, la eliminación produce un renglón cero en U. Así, $\det A = \det U = 0$. Si A es no singular, la eliminación coloca los pivotes $d_1, \ldots, d_n$ en la diagonal principal. ¡Se tiene una fórmula para el "producto de pivotes" de $\det A$! El signo depende de si el número de intercambio de renglones es par o impar:

$$\textbf{Producto de pivotes} \qquad \det A = \pm \det U = \pm d_1 d_2 \cdots d_n. \tag{1}$$

La regla nueve es la regla del producto. Yo diría que es la más sorprendente.

9. *El determinante de AB es el producto de* $\det A$ *por* $\det B$.

$$\textbf{Regla del producto} \quad |A|\,|B| = |AB| \qquad \begin{vmatrix} a & b \\ c & d \end{vmatrix}\begin{vmatrix} e & f \\ g & h \end{vmatrix} = \begin{vmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{vmatrix}.$$

Un caso particular de esta regla proporciona el determinante de A^{-1}. Debe ser $1/\det A$:

$$\det A^{-1} = \frac{1}{\det A} \quad \text{debido a que} \quad (\det A)(\det A^{-1}) = \det AA^{-1} = \det I = 1. \tag{2}$$

En el caso de 2 por 2, la regla del producto puede comprobarse pacientemente:

$$(ad - bc)(eh - fg) = (ae + bg)(cf + dh) - (af + bh)(ce + dg).$$

En el caso de n por n, se sugieren dos demostraciones posibles, ya que ésta es la regla menos evidente. En ambas demostraciones se supone que A y B son no singulares; en caso contrario, AB es singular, y la ecuación $\det AB = (\det A)(\det B)$ se comprueba fácilmente. Por la regla 8, se convierte en $0 = 0$.

i) Se demostrará que el cociente $d(A) = \det AB/\det B$ cumple las propiedades 1 a 3. Así, $d(A)$ debe ser igual a $\det A$. Por ejemplo, $d(I) = \det B/\det B = 1$; se cumple la regla 1. Si se intercambian dos renglones de A, también se intercambian los mismos renglones de AB, y el signo de d cambia según lo requiere la regla 2. Una combinación lineal en el primer renglón de A proporciona la misma combinación lineal en el primer renglón de AB. Luego la regla 3 para el determinante de AB, dividida entre la cantidad fija $\det B$, lleva a la regla 3 para el cociente $d(A)$. Así, $d(A) = \det AB/\det B$ *coincide con* $\det A$, que es nuestra fórmula del producto.

ii) Esta segunda demostración es menos elegante. Para una matriz diagonal, $\det DB = (\det D)(\det B)$ se concluye al factorizar cada d_i a partir de su renglón. Una matriz general A se reduce a D por eliminación —de A a U como de costumbre, y de U a D por eliminación hacia arriba. El determinante no cambia, excepto por una inversión de signo cuando se intercambian renglones. Los mismos pasos reducen AB a DB, con precisamente el mismo efecto sobre el determinante. Pero para DB ya se confirmó que la regla 9 es correcta.

10. *La traspuesta de A tiene el mismo determinante que A misma:* $\det A^{\mathrm{T}} = \det A$.

$$\textbf{Regla de la traspuesta} \qquad |A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & c \\ b & d \end{vmatrix} = |A^{\mathrm{T}}|.$$

De nuevo, el caso singular se trata por separado; A es singular si y sólo si A^{T} es singular, y se tiene $0 = 0$. Si A es no singular, entonces permite la factorización $PA = LDU$, y se

aplica la regla 9 para el determinante de un producto:

$$\det P \det A = \det L \det D \det U. \tag{3}$$

Al trasponer $PA = LDU$ se obtiene $A^T P^T = U^T D^T L^T$, y de nuevo por la regla 9,

$$\det A^T \det P^T = \det U^T \det D^T \det L^T. \tag{4}$$

Esto es más sencillo de lo que parece, ya que L, U, L^T, y U^T son triangulares con diagonal de 1s. Por la regla 7, su determinante es igual a 1. También, cualquier matriz diagonal es igual a su traspuesta: $D = D^T$. Sólo es necesario demostrar que $\det P = \det P^T$.

Ciertamente, $\det P$ es 1 o -1, ya que P proviene de I por intercambios de renglones. También observe que $PP^T = I$. (El 1 en el primer renglón de P corresponde al 1 en la primera columna de P^T, y le faltan los 1s en las demás columnas.) En consecuencia, $\det P \det P^T = \det I = 1$, y P y P^T deben tener el mismo determinante: ambos 1 o ambos -1.

Se concluye que los productos (3) y (4) son los mismos, y entonces $\det A = \det A^T$. Este hecho prácticamente duplica la lista de propiedades, ya que cada regla aplicada a los renglones puede aplicarse ahora a las columnas: *el determinante cambia de signo cuando se intercambian dos columnas, dos columnas iguales (o una columna de ceros) producen un determinante cero, y el determinante depende linealmente de cada columna individual.* La demostración consiste justamente en trasponer la matriz y trabajar con los renglones.

Considero que es hora de guardar silencio y declarar que la lista está completa. Sólo queda encontrar una fórmula definitiva para el determinante, y aplicarla.

Conjunto de problemas 4.2

1. Si una matriz de 4 por 4 tiene $\det A = \frac{1}{2}$, encuentre $\det(2A)$, $\det(-A)$, $\det(A^2)$, y $\det(A^{-1})$.

2. Si una matriz de 3 por 3 tiene $\det A = -1$, encuentre $\det(\frac{1}{2}A)$, $\det(-A)$, $\det(A^2)$ y $\det(A^{-1})$.

3. *Intercambio de renglones*: Sume el renglón 1 de A al renglón 2 y luego reste el renglón 2 del renglón 1. Luego sume el renglón 1 al renglón 2 y multiplique el renglón 1 por -1 para llegar a B. ¿Cuáles reglas muestran lo siguiente?

$$\det B = \begin{vmatrix} c & d \\ a & b \end{vmatrix} \quad \text{es igual a} \quad -\det A = -\begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Estas reglas pueden sustituir a la regla 2 en la definición de determinante.

4. Aplique operaciones en los renglones para obtener una triangular superior U, para calcular

$$\det \begin{bmatrix} 1 & 2 & -2 & 0 \\ 2 & 3 & -4 & 1 \\ -1 & -2 & 0 & 2 \\ 0 & 2 & 5 & 3 \end{bmatrix} \quad \text{y} \quad \det \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & -2 \end{bmatrix}.$$

Intercambie los renglones 3 y 4 de la segunda matriz, y vuelva a calcular los pivotes y el determinante.

Nota Algunos lectores ya conocerán alguna fórmula para calcular determinantes de 3 por 3. Tiene seis términos (ecuación (2) de la siguiente sección); tres son paralelos a la diagonal principal y otros tres van en dirección opuesta con signo menos. Hay una fórmula semejante para determinantes de 4 por 4, ***aunque contiene* 4! = 24 *términos*** *(no sólo ocho)*. Ni siquiera es posible estar seguro de que un signo menos va con la diagonal invertida, como se muestra en el siguiente ejercicio.

5. Cuente los intercambios de renglones para encontrar los siguientes determinantes:

$$\det \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} = +1 \quad \text{y} \quad \det \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} = -1.$$

6. Para cada n, ¿con cuántos intercambios se coloca el (renglón n, renglón $n-1, \ldots,$ renglón 1) en el orden normal (renglón $1, \ldots,$ renglón $n-1$, renglón n)? Encuentre $\det P$ para la permutación de n por n con unos en la diagonal invertida. En el problema 5 se tenía $n = 4$.

7. Encuentre el determinante de:

a) Una matriz de rango 1

$$A = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} \begin{bmatrix} 2 & -1 & 2 \end{bmatrix}.$$

b) La matriz triangular superior

$$U = \begin{bmatrix} 4 & 4 & 8 & 8 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

c) La matriz triangular inferior U^T.
d) La matriz inversa U^{-1}.
e) La "matriz triangular invertida" que resulta de intercambios de renglones,

$$M = \begin{bmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 6 \\ 0 & 1 & 2 & 2 \\ 4 & 4 & 8 & 8 \end{bmatrix}.$$

8. Demuestre cómo la regla 6 ($\det = 0$ si un renglón es cero) proviene directamente de las reglas 2 y 3.

9. Suponga que realiza dos operaciones *a la vez*, yendo de

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{a} \quad \begin{bmatrix} a - mc & b - md \\ c - \ell a & d - \ell b \end{bmatrix}.$$

Encuentre el determinante de la nueva matriz, aplicando la regla 3 o por cálculo directo.

10. Si Q es una matriz ortogonal, de modo que $Q^T Q = I$, demuestre que $\det Q$ es igual a $+1$ o -1. ¿Qué tipo de caja se forma a partir de los renglones (o las columnas) de Q?

11. Demuestre otra vez que $\det Q = 1$ o -1, usando sólo la regla del producto. Si $|\det Q| > 1$, entonces $\det Q^n$ explota. ¿Cómo se sabe que esto no puede ocurrir a Q^n?

12. Use operaciones en los renglones para justificar que el "determinante de Vandermonde" de 3 por 3 es

$$\det\begin{bmatrix}1 & a & a^2\\ 1 & b & b^2\\ 1 & c & c^2\end{bmatrix} = (b-a)(c-a)(c-b).$$

13. *a*) Una matriz simétrica sesgada cumple $K^T = -K$, como en

$$K = \begin{bmatrix}0 & a & b\\ -a & 0 & c\\ -b & -c & 0\end{bmatrix}.$$

En el caso de 3 por 3, ¿por qué $\det(-K) = (-1)^3 \det K$? Por otra parte, $\det K^T = \det K$ (siempre). Deduzca que el determinante debe ser cero.

b) Escriba una matriz simétrica sesgada de 4 por 4 con $\det K$ *diferente* de cero.

14. ¿Falso o verdadero? Proporcione una razón si es cierto, y un contraejemplo si es falso.

a) Si A y B son idénticas excepto que $b_{11} = 2a_{11}$, entonces $\det B = 2 \det A$.
b) El determinante es el producto de los pivotes.
c) Si A es invertible y B es singular, entonces $A + B$ es invertible.
d) Si A es invertible y B es singular, entonces AB es singular.
e) El determinante de $AB - BA$ es cero.

15. Si la suma de cada renglón de A es cero, demuestre que $\det A = 0$. Si la suma de cada renglón es 1, demuestre que $\det(A - I) = 0$. Demuestre con un ejemplo que esto no implica que $\det A = I$.

16. Encuentre los siguientes determinantes de 4 por 4 por eliminación gaussiana:

$$\det\begin{bmatrix}11 & 12 & 13 & 14\\ 21 & 22 & 23 & 24\\ 31 & 32 & 33 & 34\\ 41 & 42 & 43 & 44\end{bmatrix} \quad \text{y} \quad \det\begin{bmatrix}1 & t & t^2 & t^3\\ t & 1 & t & t^2\\ t^2 & t & 1 & t\\ t^3 & t^2 & t & 1\end{bmatrix}.$$

17. Encuentre los determinantes de

$$A = \begin{bmatrix}4 & 2\\ 1 & 3\end{bmatrix}, \quad A^{-1} = \frac{1}{10}\begin{bmatrix}3 & -2\\ -1 & 4\end{bmatrix}, \quad A - \lambda I = \begin{bmatrix}4-\lambda & 2\\ 1 & 3-\lambda\end{bmatrix}.$$

¿Para qué valores de λ se cumple que $A - \lambda I$ es una matriz singular?

18. Evalúe $\det A$ reduciendo la matriz a forma triangular (reglas 5 y 7).

$$A = \begin{bmatrix}1 & 1 & 3\\ 0 & 4 & 6\\ 1 & 5 & 8\end{bmatrix}, \quad B = \begin{bmatrix}1 & 1 & 3\\ 0 & 4 & 6\\ 0 & 0 & 1\end{bmatrix}, \quad C = \begin{bmatrix}1 & 1 & 3\\ 0 & 4 & 6\\ 1 & 5 & 9\end{bmatrix}.$$

¿Cuáles son los determinantes de B, C, AB, A^TA, y C^T?

19. Suponga que $CD = -DC$, y encuentre la falla en el siguiente razonamiento: Al tomar determinantes se obtiene $(\det C)(\det D) = -(\det D)(\det C)$, de modo que $\det C = 0$ o $\det D = 0$. Así, $CD = -DC$ sólo es posible si C o D es singular.

20. ¿El determinante de las siguientes matrices es igual a 0, 1, 2, o 3?

$$A = \begin{bmatrix}0 & 0 & 1\\ 1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix} \quad B = \begin{bmatrix}0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 0\end{bmatrix} \quad C = \begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}.$$

21. Parece que la inversa de una matriz de 2 por 2 tiene determinante = 1:

$$\det A^{-1} = \det \frac{1}{ad-bc}\begin{bmatrix}d & -b\\ -c & a\end{bmatrix} = \frac{ad-bc}{ad-bc} = 1.$$

¿Qué está mal en este cálculo? ¿Por qué es correcto $\det A^{-1}$?

En los problemas 22 a 28 se usan reglas para calcular determinantes específicos.

22. Reduzca A a U, y encuentre $\det A$ = producto de los pivotes:

$$A = \begin{bmatrix}1 & 1 & 1\\ 1 & 2 & 2\\ 1 & 2 & 3\end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix}1 & 2 & 3\\ 2 & 2 & 3\\ 3 & 3 & 3\end{bmatrix}.$$

23. Aplique operaciones en los renglones para obtener una triangular superior U, y calcule

$$\det\begin{bmatrix}1 & 2 & 3 & 0\\ 2 & 6 & 6 & 1\\ -1 & 0 & 0 & 3\\ 0 & 2 & 0 & 7\end{bmatrix} \quad \text{y} \quad \det\begin{bmatrix}2 & 1 & 1 & 1\\ 1 & 2 & 1 & 1\\ 1 & 1 & 2 & 1\\ 1 & 1 & 1 & 2\end{bmatrix}.$$

24. Aplique operaciones en los renglones para simplificar y calcular los siguientes determinantes:

$$\det\begin{bmatrix}101 & 201 & 301\\ 102 & 202 & 302\\ 103 & 203 & 303\end{bmatrix} \quad \text{y} \quad \det\begin{bmatrix}1 & t & t^2\\ t & 1 & t\\ t^2 & t & 1\end{bmatrix}.$$

25. La eliminación reduce A a U. Luego, $A = LU$:

$$A = \begin{bmatrix}3 & 3 & 4\\ 6 & 8 & 7\\ -3 & 5 & -9\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ -1 & 4 & 1\end{bmatrix}\begin{bmatrix}3 & 3 & 4\\ 0 & 2 & -1\\ 0 & 0 & -1\end{bmatrix} = LU.$$

Encuentre los determinantes de L, U, A, $U^{-1}L^{-1}$, y $U^{-1}L^{-1}A$.

26. Si a_{ij} es i multiplicado por j, demuestre que $\det A = 0$. (Con la excepción cuando $A = [1]$.)

27. Si a_{ij} es $i + j$, demuestre que $\det A = 0$. (Con la excepción cuando $n = 1$ o 2.)

28. Calcule los determinantes de las siguientes matrices, mediante operaciones en los renglones:

$$A = \begin{bmatrix}0 & a & 0\\ 0 & 0 & b\\ c & 0 & 0\end{bmatrix}, \quad B = \begin{bmatrix}0 & a & 0 & 0\\ 0 & 0 & b & 0\\ 0 & 0 & 0 & c\\ d & 0 & 0 & 0\end{bmatrix}, \quad \text{y} \quad C = \begin{bmatrix}a & a & a\\ a & b & b\\ a & b & c\end{bmatrix}.$$

29. ¿Cuál es el error en la siguiente demostración de que las matrices proyección cumplen $\det P = 1$?

$$P = A(A^TA)^{-1}A^T \quad \text{de modo que} \quad |P| = |A|\frac{1}{|A^T|\,|A|}|A^T| = 1.$$

30. (Problema de cálculo) Demuestre que las derivadas parciales de ln(det A) proporcionan A^{-1}:

$$f(a, b, c, d) = \ln(ad - bc) \quad \text{conduce a} \quad \begin{bmatrix} \partial f/\partial a & \partial f/\partial c \\ \partial f/\partial b & \partial f/\partial d \end{bmatrix} = A^{-1}.$$

31. (MATLAB) La matriz de Hilbert **hilb**(n) tiene el elemento i, j igual a $1/(i + j - 1)$. Imprima los determinantes de **hilb**(1), **hilb**(2), . . . , **hilb**(10). ¡Es difícil trabajar con las matrices de Hilbert! ¿Cuáles son los pivotes?

32. (MATLAB) ¿Cuál es un determinante típico (experimentalmente) de **rand**(n) y **randn**(n) para n = 50, 100, 200, 400? (¿Y qué significa "Inf" en MATLAB?)

33. Use MATLAB para encontrar el determinante más grande de una matriz de 4 por 4 de 0s y 1s.

34. Si se sabe que det A = 6, ¿cuál es el determinante de B?

$$\det A = \begin{vmatrix} \text{renglón 1} \\ \text{renglón 2} \\ \text{renglón 3} \end{vmatrix} = 6 \qquad \det B = \begin{vmatrix} \text{renglón 1 + renglón 2} \\ \text{renglón 2 + renglón 3} \\ \text{renglón 3 + renglón 1} \end{vmatrix} = __$$

35. Suponga que la matriz M de 4 por 4 tiene cuatro renglones iguales, todos conteniendo a, b, c, d. Se sabe que det(M) = 0. El problema es encontrar det($I + M$) por cualquier método:

$$\det(I + M) = \begin{vmatrix} 1+a & b & c & d \\ a & 1+b & c & d \\ a & b & 1+c & d \\ a & b & c & 1+d \end{vmatrix}.$$

Crédito parcial si encuentra este determinante cuando $a = b = c = d = 1$. Eliminado de inmediato si afirma que det($I + M$) = det I + det M.

4.3 FÓRMULAS PARA EL DETERMINANTE

La primera fórmula ya apareció. Las operaciones en los renglones producen los pivotes en D:

4A Si A es invertible, entonces $PA = LDU$ y det $P = \pm 1$. La regla del producto da

$$\det A = \pm \det L \det D \det U = \pm\ \textbf{(producto de los pivotes).} \qquad (1)$$

El signo ± 1 depende de si el número de intercambios de renglones es par o impar. Los factores triangulares tienen det L = det U = 1 y det $D = d_1 \cdots d_n$.

En el caso de 2 por 2, la factorización estándar LDU es

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ c/a & 1 \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & (ad - bc)/a \end{bmatrix} \begin{bmatrix} 1 & b/a \\ 0 & 1 \end{bmatrix}.$$

El producto de los pivotes es $ad - bc$. Este es el determinante de la matriz diagonal D. Si el primer paso es un intercambio de renglones, entonces los pivotes son c y $(-\det A)/c$.

Ejemplo 1 Los pivotes de la matriz $-1, 2, -1$ en segundas diferencias son 2/1, 3/2, . . . en D:

$$\begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & \cdot & \\ & & \cdot & \cdot & -1 \\ & & & -1 & 2 \end{bmatrix} = LDU = L \begin{bmatrix} 2 & & & & \\ & 3/2 & & & \\ & & 4/3 & & \\ & & & \cdot & \\ & & & & (n+1)/n \end{bmatrix} U.$$

Su determinante es el producto de los pivotes. Todos los números 2, . . ., n se cancelan:

$$\det A = 2\left(\frac{3}{2}\right)\left(\frac{4}{3}\right)\cdots\left(\frac{n+1}{n}\right) = n + 1.$$

MATLAB calcula el determinante a partir de los pivotes. Sin embargo, concentrar toda la información en los pivotes hace imposible imaginarse cómo el cambio en un elemento puede afectar al determinante. Se desea encontrar una expresión explícita para el determinante en términos de los n^2 elementos.

Para $n = 2$, se demostrará que $ad - bc$ es correcto. Para $n = 3$, la fórmula del determinante es bastante bien conocida (tiene seis términos):

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{matrix} +a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ -a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \end{matrix} \qquad (2)$$

Nuestro objetivo es deducir estas fórmulas directamente a partir de las propiedades definitorias 1 a 3 de det A. Si es posible manejar $n = 2$ y $n = 3$ de manera organizada, podrá verse el patrón.

Para empezar, cada renglón puede descomponerse en vectores en las direcciones de coordenadas:

$$\begin{bmatrix} a & b \end{bmatrix} = \begin{bmatrix} a & 0 \end{bmatrix} + \begin{bmatrix} 0 & b \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} c & d \end{bmatrix} = \begin{bmatrix} c & 0 \end{bmatrix} + \begin{bmatrix} 0 & d \end{bmatrix}.$$

Luego se aplica la propiedad de linealidad, primero al renglón 1 y luego al renglón 2:

Separar en $n^n = 2^2$ determinantes fáciles

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & 0 \\ c & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & d \end{vmatrix}$$
$$= \begin{vmatrix} a & 0 \\ c & 0 \end{vmatrix} + \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} + \begin{vmatrix} 0 & b \\ 0 & d \end{vmatrix}. \qquad (3)$$

Cada renglón consta de n direcciones de coordenadas, por lo que su desarrollo contiene n^2 términos. La mayor parte de tales términos (todos menos $n! = n$ factorial) son automáticamente cero. Cuando dos renglones están en la misma dirección de coordenadas, uno es múltiplo del otro, y

$$\begin{vmatrix} a & 0 \\ c & 0 \end{vmatrix} = 0, \qquad \begin{vmatrix} 0 & b \\ 0 & d \end{vmatrix} = 0.$$

Se presta atención *sólo cuando los renglones apuntan en distintas direcciones*. ***Los términos diferentes de cero deben estar en columnas distintas.*** Suponga que el primer renglón tiene un elemento diferente de cero en la columna α, el segundo, un elemento diferente de cero en la columna β, y que finalmente el renglón n-ésimo tiene un elemento diferente de cero en la columna v. Todos los números de columna $\alpha, \beta, \ldots, v$ son diferentes. Son un reordenamiento, o ***permutación***, de los números 1, 2, . . . , n. El caso de 3 por 3 pro-

duce $3! = 6$ determinantes:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & & \\ & a_{22} & \\ & & a_{33} \end{vmatrix} + \begin{vmatrix} & a_{12} & \\ & & a_{23} \\ a_{31} & & \end{vmatrix} + \begin{vmatrix} & & a_{13} \\ a_{21} & & \\ & a_{32} & \end{vmatrix}$$

$$+ \begin{vmatrix} a_{11} & & \\ & & a_{23} \\ & a_{32} & \end{vmatrix} + \begin{vmatrix} & a_{12} & \\ a_{21} & & \\ & & a_{33} \end{vmatrix} + \begin{vmatrix} & & a_{13} \\ & a_{22} & \\ a_{31} & & \end{vmatrix}. \quad (4)$$

Excepto estos $n!$ determinantes, todos los demás son cero porque la columna se repite. (Para la primera columna α hay n opciones, para β hay $n - 1$ opciones restantes, y por último, queda una opción para la última columna v. Cada vez se utiliza una sola columna, cuando se "serpentean" los renglones de la matriz). En otras palabras, ***hay n! formas de permutar los números* 1, 2, . . . , *n***. Los números de columna proporcionan las permutaciones:

Números de columna $(\alpha, \beta, v) = (1, 2, 3), (\mathbf{2, 3, 1}), (3, 1, 2), (1, 3, 2), (2, 1, 3), (3, 2, 1).$

Hay $3! = 6$ permutaciones de (1, 2, 3); la primera es la identidad.

El determinante de A se reduce ahora a seis determinantes por separado y mucho más sencillos. Al factorizar a_{ij}, para cada una de las seis permutaciones hay un término:

$$\det A = a_{11}a_{22}a_{33}\begin{vmatrix} 1 & & \\ & 1 & \\ & & 1 \end{vmatrix} + a_{12}a_{23}a_{31}\begin{vmatrix} & 1 & \\ & & 1 \\ 1 & & \end{vmatrix} + a_{13}a_{21}a_{32}\begin{vmatrix} & & 1 \\ 1 & & \\ & 1 & \end{vmatrix}$$

$$+ a_{11}a_{23}a_{32}\begin{vmatrix} 1 & & \\ & & 1 \\ & 1 & \end{vmatrix} + a_{12}a_{21}a_{33}\begin{vmatrix} & 1 & \\ 1 & & \\ & & 1 \end{vmatrix} + a_{13}a_{22}a_{31}\begin{vmatrix} & & 1 \\ & 1 & \\ 1 & & \end{vmatrix}. \quad (5)$$

Cada término es un producto de $n = 3$ elementos a_{ij}, donde *cada renglón y cada columna están representados una vez*. Si el orden de las columnas es $(\alpha, \ldots, v)$, ese término es el producto de $a_{1\alpha} \cdots a_{nv}$ por el determinante de una matriz permutación P. El determinante de toda la matriz es la suma de estos $n!$ términos, y ***esa sumatoria es la fórmula explícita que se está buscando***:

Gran fórmula $$\det A = \sum_{\text{todas las } Ps} (a_{1\alpha}a_{2\beta} \cdots a_{nv}) \det P. \quad (6)$$

Para una matriz de n por n, esta sumatoria se toma sobre todas las $n!$ permutaciones $(\alpha, \ldots, v)$ de los números $(1, \ldots, n)$. La permutación proporciona los números de columna a medida que se desciende por la matriz. Los 1s aparecen en P en los mismos sitios en que las as aparecían en A.

Queda por encontrar el determinante de P. Los intercambios de renglones transforman P en la matriz identidad, y cada intercambio cambia el signo del determinante:

$\det P = +1$ o -1 ***para un número par* o *impar de intercambios de renglones*.**

$$(1, 3, 2) \ \text{es impar, de modo que} \begin{vmatrix} 1 & & \\ & & 1 \\ & 1 & \end{vmatrix} = -1 \quad (3, 1, 2) \ \text{es par, de modo que} \begin{vmatrix} & & 1 \\ 1 & & \\ & 1 & \end{vmatrix} = 1$$

(1, 3, 2) requiere un intercambio y (3, 1, 2) requiere dos intercambios para recuperar (1, 2, 3). Estos son dos de los seis signos $\pm$. Para $n = 2$, sólo se tiene (1, 2) y (2, 1):

$$\det A = a_{11}a_{22} \det\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + a_{12}a_{21} \det\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21} \ (\text{o } ad - bc).$$

Nadie puede afirmar que la gran fórmula (6) es particularmente simple. Sin embargo, es posible darse cuenta de por qué tiene las propiedades 1 a 3. Para $A = I$, todo producto de los a_{ij} es cero, excepto por la sucesión de columnas $(1, 2, \ldots, n)$. Este término proporciona $\det I = 1$. La propiedad 2 se comprobará en la siguiente sección, porque aquí se tiene más interés en la propiedad 3: el determinante debe depender linealmente del primer renglón $a_{11}, a_{12}, \ldots, a_{1n}$.

Considere a todos los términos $a_{1\alpha}a_{2\beta} \cdots a_{nv}$ que implican a a_{11}. La primera columna es $\alpha = 1$. Esto deja alguna permutación $(\beta, \ldots, v)$ de las columnas restantes $(2, \ldots, n)$. Todos estos términos se agrupan como $a_{11}C_{11}$, donde el coeficiente de a_{11} es un determinante más pequeño, *al que se han retirado el renglón* 1 *y la columna* 1:

Cofactor de a_{11} $$C_{11} = \sum (a_{2\beta} \cdots a_{nv}) \det P = \det \ (\text{submatriz de A}) \quad (7)$$

De manera semejante, el elemento a_{12} se ha multiplicado por algún determinante más pequeño C_{12}. Al agrupar todos los términos que empiezan con el mismo a_{1j}, la fórmula (6) se convierte en

Cofactores a lo largo del renglón 1 $$\det A = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}. \quad (8)$$

Esto demuestra que $\det A$ depende linealmente de los elementos $a_{11}, \ldots, a_{1n}$ del primer renglón.

Ejemplo 2 Para una matriz de 3 por 3, con esta manera de agrupar términos se obtiene

$$\det A = a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{12}(a_{23}a_{31} - a_{21}a_{33}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}). \quad (9)$$

Los *cofactores* C_{11}, C_{12}, C_{13} son los determinantes de 2 por 2 entre paréntesis.

Desarrollo de det *A* por cofactores

Se busca otra fórmula más para el determinante. Si esto significa partir de cero, sería demasiado. Sin embargo, *la fórmula ya está descubierta: es* (8), *y la única cuestión es identificar los cofactores* C_{1j} *que multiplican a* a_{1j}.

Se sabe que C_{1j} depende de los renglones $2, \ldots, n$. El renglón 1 ya está tomado en cuenta por a_{1j}. Además, a_{1j} también toma en cuenta la columna j-ésima, de modo que su cofactor C_{1j} debe depender por completo de *las otras columnas*. Ningún renglón ni columna puede usarse dos veces en el mismo término. Lo que realmente se está haciendo es separar el determinante en la siguiente suma:

Separación en cofactores $$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & & \\ & a_{22} & a_{23} \\ & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} & a_{12} & \\ a_{21} & & a_{23} \\ a_{31} & & a_{33} \end{vmatrix} + \begin{vmatrix} & & a_{13} \\ a_{21} & a_{22} & \\ a_{31} & a_{32} & \end{vmatrix}.$$

Para un determinante de orden n, esta separación origina n determinantes más pequeños (***menores***) de orden $n - 1$; se puede ver las tres submatrices de 2 por 2. La submatriz M_{1j} se forma *quitando el renglón* 1 *y la columna* j. Su determinante se multiplica por a_{1j}, así como por un signo más o menos. Estos signos alternan como en $\det M_{11}$, $-\det M_{12}$, $\det M_{13}$:

$$\textbf{Cofactores del renglón 1} \qquad C_{1j} = (-1)^{1+j} \det M_{1j}.$$

El segundo cofactor C_{12} es $a_{23}a_{31} - a_{21}a_{33}$, que es $\det M_{12}$ multiplicado por -1. Esta misma técnica funciona en toda matriz de n por n. La separación anterior confirma que C_{11} es el determinante del ángulo inferior derecho M_{11}.

Hay un desarrollo semejante en cualquier otro renglón, por ejemplo i. Puede demostrarse intercambiando el renglón i con el renglón 1. *Recuerde borrar el renglón i y la columna j de A para M_{ij}:*

4B El determinante de A es una combinación de cualquier renglón i multiplicado por sus cofactores:

$$\textbf{det } A \textbf{ por cofactores} \qquad \det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}. \tag{10}$$

El cofactor C_{ij} es el determinante de M_{ij} con el signo correcto:

$$\textbf{borrar el renglón } i \textbf{ y la columna } j \qquad C_{ij} = (-1)^{i+j} \det M_{ij}. \tag{11}$$

Estas fórmulas expresan $\det A$ como una combinación de determinantes de orden $n - 1$. *Hubiera sido posible definir el determinante por inducción sobre n.* Una matriz de 1 por 1 tiene $\det A = a_{11}$, y así la ecuación (10) define los determinantes de las matrices de 2 por 2, de 3 por 3, y de n por n. Aquí se prefirió definir un determinante según sus propiedades, que son mucho más simples de explicar. La fórmula explícita (6) y la fórmula de cofactores (10) se concluyeron directamente, a partir de esas propiedades.

Hay una consecuencia más de $\det A = \det A^{\mathrm{T}}$. Es posible desarrollar por cofactores de una *columna* de A, que es un renglón de A^{T}. Siguiendo por la columna j de A,

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}. \tag{12}$$

Ejemplo 3 La matriz en segundas diferencias A_4 de 4 por 4 sólo tiene dos elementos diferentes de cero en el renglón 1:

$$\textbf{Usar cofactores} \qquad A_4 = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}.$$

C_{11} proviene al borrar el renglón 1 y la columna 1, quedando así el patrón $-1, 2, -1$:

$$C_{11} = \det A_3 = \det \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}.$$

Para $a_{12} = -1$, la columna eliminada es la 2, y se requiere su cofactor C_{12}:

$$C_{12} = (-1)^{1+2} \det \begin{bmatrix} -1 & -1 & 0 \\ 0 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = +\det \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \det A_2.$$

Así se queda con el determinante de 2 por 2. En total, el renglón 1 ha producido $2C_{11} - C_{12}$:

$$\det A_4 = 2(\det A_3) - \det A_2 = 2(4) - 3 = 5$$

La misma idea es válida para A_5 y A_6, así como para toda A_n:

$$\textbf{Recurrencia por cofactores} \qquad \det A_n = 2(\det A_{n-1}) - \det A_{n-2}. \tag{13}$$

Lo anterior proporciona el determinante de matrices cada vez más grandes. En cada paso, el determinante de A_n es $n + 1$, a partir de los determinantes previos n y $n - 1$:

$$\textbf{matriz } \mathbf{-1, 2, -1} \qquad \det A_n = 2(n) - (n - 1) = n + 1.$$

La respuesta $n + 1$ coincide con el producto de pivotes al inicio de esta sección.

Conjunto de problemas 4.3

1. Para las siguientes matrices, encuentre el único elemento diferente de cero en la gran fórmula (6):

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0 & 0 & 1 & 2 \\ 0 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Sólo hay una forma de elegir cuatro elementos diferentes de cero de renglones distintos y columnas distintas. Al decidir par o impar, calcule $\det A$ y $\det B$.

2. Desarrolle estos determinantes en cofactores en el primer renglón. Encuentre los cofactores (incluyen los signos $(-1)^{i+j}$) y los determinantes de A y B.

3. *¿Falso o verdadero?*
a) El determinante de $S^{-1}AS$ es igual al determinante de A.
b) Si $\det A = 0$, entonces por lo menos uno de los cofactores debe ser cero.
c) Una matriz cuyos elementos son 0s y 1s tiene determinante 1, 0, o -1.

4. *a)* Encuentre la factorización LU, los pivotes, y el determinante de la matriz de 4 por 4 cuyos elementos son a_{ij} = el menor de i y j. (Escriba la matriz.)
b) Encuentre el determinante si a_{ij} = el menor de n_i y n_j, donde $n_1 = 2$, $n_2 = 6$, $n_3 = 8$, $n_4 = 10$. ¿Puede proporcionar una regla general para cualquier $n_1 \le n_2 \le n_3 \le n_4$?

5. Sea F_n el determinante de la matriz tridiagonal $1, 1, -1$ (de n por n):

$$F_n = \det \begin{bmatrix} 1 & -1 & & & \\ 1 & 1 & -1 & & \\ & 1 & 1 & -1 & \\ & & \cdot & \cdot & \cdot \\ & & & 1 & 1 \end{bmatrix}.$$

Desarrolle por cofactores a lo largo del renglón 1, para demostrar que $F_n = F_{n-1} + F_{n-2}$. Esto lleva a la *sucesión de Fibonacci* 1, 2, 3, 5, 8, 13, ... para los determinantes.

6. Suponga que A_n es la matriz tridiagonal n por n con 1s en las tres diagonales:

$$A_1 = [1], \qquad A_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad A_3 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad \ldots$$

Sea D_n el determinante de A_n; se requiere encontrarlo.

a) Desarrolle por cofactores a lo largo del primer renglón para demostrar que $D_n = D_{n-1} - D_{n-2}$.
b) Empiece con $D_1 = 1$ y $D_2 = 0$, para encontrar $D_3, D_4, \ldots, D_8$. Observe la manera en que estos números se repiten (¿con qué periodo?), y encuentre D_{1000}.

7. *a*) Evalúe el siguiente determinante por cofactores del renglón 1:

$$\begin{vmatrix} 4 & 4 & 4 & 4 \\ 1 & 2 & 0 & 1 \\ 2 & 0 & 1 & 2 \\ 1 & 1 & 0 & 2 \end{vmatrix}.$$

b) Compruebe restando la columna 1 de las demás columnas, y vuelva a calcular.

8. Calcule los determinantes de A_2, A_3, A_4. ¿Puede pronosticar A_n?

$$A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}.$$

Use operaciones en los renglones para producir ceros, o use los cofactores del renglón 1.

9. ¿Cuántas multiplicaciones se requieren para encontrar un determinante de n por n a partir de

a) la gran fórmula (6)?
b) la fórmula de cofactores (10), construyendo a partir del conteo para $n - 1$?
c) la fórmula del producto de pivotes (incluyendo los pasos de eliminación)?

10. En una matriz de 5 por 5, ¿un signo + o un signo − con $a_{15}a_{24}a_{33}a_{42}a_{51}$ invierte la diagonal? En otras palabras, $P = (5, 4, 3, 2, 1)$ ¿es par o impar? El patrón de tablero de ajedrez de signos ± para cofactores *no* proporciona det P.

11. Si A es de m por n y B es de n por m, explique por qué

$$\det \begin{bmatrix} 0 & A \\ -B & I \end{bmatrix} = \det AB. \qquad \left(\begin{array}{c} \textit{Sugerencia}\text{: posmultiplicar} \\ \text{por la derecha por)} \end{array} \begin{bmatrix} I & 0 \\ B & I \end{bmatrix}. \right)$$

Proporcione un ejemplo con $m < n$, y un ejemplo con $m > n$. ¿Por qué su segundo ejemplo automáticamente tiene det $AB = 0$?

12. Suponga que la matriz A es fija, excepto que a_{11} varía desde $-\infty$ hasta $+\infty$. Proporcione ejemplos en los cuales det A siempre es cero o nunca es cero. Luego, a partir del desarrollo por cofactores (8), demuestre que en caso contrario det $A = 0$ para exactamente *un valor* de a_{11}.

En los problemas 13 a 23, use la gran fórmula con $n!$ términos: $|A| = \sum \pm a_{1\alpha}a_{2\beta} \cdots a_{n\nu}$.

13. Calcule los determinantes de A, B, C, a partir de seis términos. ¿Sus renglones son independientes?

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 3 & 2 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 4 & 4 \\ 5 & 6 & 7 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

14. Calcule los determinantes de A, B, C. ¿Sus columnas son independientes?

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad C = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}.$$

15. Demuestre que det $A = 0$, sin importar los cinco elementos diferentes de cero indicados por las xs:

$$A = \begin{bmatrix} x & x & x \\ 0 & 0 & x \\ 0 & 0 & x \end{bmatrix}. \qquad (\text{¿Cuál es el rango de } A\text{?})$$

16. Este problema muestra en dos formas que det $A = 0$ (las xs son números cualesquiera):

$$A = \begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & x & x \end{bmatrix}. \qquad \begin{array}{l} \text{matriz de 5 por 5} \\ \text{matriz cero de 3 por 3} \\ \text{Siempre es singular} \end{array}$$

a) ¿Cómo se sabe que los renglones son linealmente dependientes?
b) Explique por qué todos los 120 términos son cero en la gran fórmula para det A.

17. Encuentre dos formas para elegir elementos diferentes de cero, a partir de cuatro renglones y columnas diferentes:

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 3 & 4 & 5 \\ 5 & 4 & 0 & 3 \\ 2 & 0 & 0 & 1 \end{bmatrix}. \ (B \text{ tiene los mismos ceros que } A).$$

det A, ¿es igual a $1 + 1$ o $1 - 1$ o $-1 - 1$? ¿A qué es igual det B?

18. Coloque el menor número de ceros en una matriz de 4 por 4 que garantice det $A = 0$. Coloque tantos ceros como sea posible permitiendo a la vez que det $A \neq 0$.

19. *a*) Si $a_{11} = a_{22} = a_{33} = 0$, ¿cuántos de los seis términos en det A son cero?
b) Si $a_{11} = a_{22} = a_{33} = a_{44} = 0$, ¿cuántos de los 24 productos $a_{1j}a_{2k}a_{3\ell}a_{4m}$ en det A es seguro que son cero?

20. ¿Cuántas matrices permutación de 5 por 5 tienen det $P = +1$? Estas son permutaciones pares. Encuentre una que requiera cuatro intercambios para llegar a la matriz identidad.

21. Si det $A \neq 0$, por lo menos uno de los $n!$ términos en la gran fórmula (6) no es cero. Deduzca que algún ordenamiento de los renglones de A no deja ceros en la diagonal. (No use P de la eliminación; esa PA puede tener ceros en la diagonal.)

22. Demuestre que 4 es el determinante más grande para una matriz de 3 por 3 de 1s positivos y 1s negativos.

23. ¿Cuántas permutaciones de (1, 2, 3, 4) son pares, y cuáles son? Crédito extra: ¿Cuáles son todos los posibles determinantes de 4 por 4 de $I + P_{\text{par}}$?

En los problemas 24 a 33 se usan cofactores $C_{ij} = (-1)^{i+j} \det M_{ij}$. Elimine el renglón i, columna j.

24. Encuentre cofactores, y luego trasponga. ¡Multiplique C_A^{T} y C_B^{T} por A y B!

$$A = \begin{bmatrix} 2 & 1 \\ 3 & 6 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 0 & 0 \end{bmatrix}.$$

25. Encuentre la matriz por cofactores C, y compare AC^T con A^{-1}:

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \qquad A^{-1} = \frac{1}{4}\begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

26. La matriz B_n es la matriz A_n $-1, 2, -1$, excepto que $b_{11} = 1$ en vez de $a_{11} = 2$. Use cofactores del *último* renglón de B_4, para demostrar que $|B_4| = 2|B_3| - |B_2| = 1$:

$$B_4 = \begin{bmatrix} 1 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 2 \end{bmatrix} \qquad B_3 = \begin{bmatrix} 1 & -1 & \\ -1 & 2 & -1 \\ & -1 & 2 \end{bmatrix}.$$

La recurrencia $|B_n| = 2|B_{n-1}| - |B_{n-2}|$ es la misma que para las As. La diferencia es en los valores iniciales 1, 1, 1 para $n = 1, 2, 3$. *¿Cuáles son los pivotes?*

27. B_n sigue siendo igual a A_n, excepto por $b_{11} = 1$. Así, use linealidad en el primer renglón, donde [1 −1 0] es igual a [2 −1 0] menos [1 0 0]:

$$|B_n| = \begin{vmatrix} 1 & -1 & & 0 \\ -1 & & & \\ & & A_{n-1} & \\ 0 & & & \end{vmatrix} = \begin{vmatrix} 2 & -1 & & 0 \\ -1 & & & \\ & & A_{n-1} & \\ 0 & & & \end{vmatrix} - \begin{vmatrix} 1 & 0 & & 0 \\ -1 & & & \\ & & A_{n-1} & \\ 0 & & & \end{vmatrix}.$$

La linealidad en el renglón 1 proporciona $|B_n| = |A_n| - |A_{n-1}| =$ ___.

28. El determinante C_n de n por n tiene 1s arriba y abajo de la diagonal principal:

$$C_1 = |0| \quad C_2 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} \quad C_3 = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} \quad C_4 = \begin{vmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{vmatrix}.$$

a) ¿Cuáles son los determinantes de C_1, C_2, C_3, C_4?
b) Por cofactores, encuentre la relación entre C_n y C_{n-1} y C_{n-2}. Encuentre C_{10}.

29. En el problema 28 se tienen 1s justo arriba y abajo de la diagonal principal. Descendiendo en la matriz, ¿qué orden de las columnas (en caso de haber alguno) proporciona todos los 1s? Explique por qué esa permutación es *par* para $n = 4, 8, 12, \ldots$, e *impar* para $n = 2, 6, 10, \ldots$

$$C_n = 0\ (n \text{ impar}) \qquad C_n = 1\ (n = 4, 8, \ldots) \qquad C_n = -1\ (n = 2, 6, \ldots).$$

30. Explique por qué el siguiente determinante de Vandermonde contiene a x^3 pero no a x^4 o x^5:

$$V_4 = \det \begin{bmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & x & x^2 & x^3 \end{bmatrix}.$$

El determinante es cero en $x =$ ___, ___, y ___. El cofactor de x^3 es $V_3 = (b-a)(c-a)(c-b)$. Así, $V_4 = (x-a)(x-b)(x-c)V_3$.

31. Calcule los determinantes S_1, S_2, S_3 de las siguientes matrices tridiagonales:

$$S_1 = |3| \qquad S_2 = \begin{vmatrix} 3 & 1 \\ 1 & 3 \end{vmatrix} \qquad S_3 = \begin{vmatrix} 3 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 3 \end{vmatrix}.$$

Haga una conjetura de Fibonacci para S_4, y compruebe que tiene razón.

32. Los cofactores de estas matrices 1, 3 1 proporcionan $S_n = 3S_{n-1} - S_{n-2}$. Desafío: *Demuestre que S_n es el número de Fibonacci F_{2n+2} al demostrar $F_{2n+2} = 3F_{2n} - F_{2n-2}$.* Siga utilizando la regla de Fibonacci $F_k = F_{k-1} + F_{k-2}$.

33. Cambie 3 por 2 en la esquina superior izquierda de las matrices en el problema 32. ¿Por qué se resta S_{n-1} del determinante S_n? Demuestre que los determinantes se convierten en los números de Fibonacci 2, 5, 13 (siempre F_{2n+1}).

Los problemas 34 a 36 son sobre matrices en bloque y determinantes en bloque.

34. Con bloques de 2 por 2, ¡no siempre es posible usar determinantes en bloque!

$$\begin{vmatrix} A & B \\ 0 & D \end{vmatrix} = |A|\,|D| \qquad \text{pero} \qquad \begin{vmatrix} A & B \\ C & D \end{vmatrix} \neq |A|\,|D| - |C|\,|B|.$$

a) ¿Por qué es cierta la primera afirmación? De alguna manera, B no entra.
b) Demuestre con un ejemplo que la igualdad falla (como se muestra) cuando entra C.
c) Demuestre con un ejemplo que la respuesta $\det(AD - CB)$ también es errónea.

35. Con multiplicación por bloques, $A = LU$ tiene $A_k = L_kU_k$ en la esquina superior izquierda:

$$A = \begin{bmatrix} A_k & * \\ * & * \end{bmatrix} = \begin{bmatrix} L_k & 0 \\ * & * \end{bmatrix} \begin{bmatrix} U_k & * \\ 0 & * \end{bmatrix}.$$

a) Suponga que los tres primeros pivotes de A son 2, 3, −1. ¿Cuáles son los determinantes de L_1, L_2, L_3 (con 1s en la diagonal), U_1, U_2, U_3, y A_1, A_2, A_3?
b) Si A_1, A_2, A_3 tienen determinantes 5, 6, 7, encuentre los tres pivotes.

36. En la eliminación por bloque se resta CA^{-1} multiplicado por el primer renglón $[A\ B]$ del segundo renglón $[C\ D]$. Así, en la esquina queda el *complemento de Schur* $D - CA^{-1}B$:

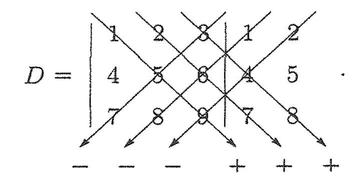
$$\begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix}.$$

Tome determinantes de estas matrices para demostrar reglas correctas para bloques cuadrados:

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = \underset{\text{si } A^{-1} \text{ existe}}{|A|\,|D - CA^{-1}B|} = \underset{\text{si } AC = CA}{|AD - CB|}.$$

37. Un determinante de 3 por 3 tiene tres productos "abajo a la derecha" y tres "abajo a la izquierda" con signos menos. Calcule los seis términos de la figura para encontrar D.

Luego explique sin determinantes por qué la siguiente matriz es invertible o no:



38. Para A_4 en el problema 6, cinco de los $4! = 24$ términos en la gran fórmula (6) son diferentes de cero. Encuentre estos cinco términos para demostrar que $D_4 = -1$.

39. Para la matriz tridiagonal de 4 por 4 (elementos $-1, 2, -1$), encuentre los cinco términos en la gran fórmula con los que se obtiene $\det A = 16 - 4 - 4 - 4 + 1$.

40. Encuentre el determinante de la siguiente matriz cíclica P por cofactores del renglón 1. ¿Con cuántos intercambios se reordena 4, 1, 2, 3 en 1, 2, 3, 4? es ¿$|P^2| = +1$ o -1?

$$P = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad P^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}.$$

41. $A = 2 * \text{eye}(n) - \text{diag}(\text{ones}(n-1, 1), 1) - \text{diag}(\text{ones}(n-1, 1), -1)$ es la matriz $-1, 2, -1$. Cambie $A(1, 1)$ a 1 de modo que $\det A = 1$. Pronostique los elementos de A^{-1} con base en $n = 3$, y pruebe su conjetura para $n = 4$.

42. (MATLAB) Las matrices $-1, 2, -1$ tienen determinante $n + 1$. Calcule $(n + 1)A^{-1}$ para $n = 3$ y 4, y compruebe su conjetura para $n = 5$. (Las inversas de las matrices tridiagonales tienen la forma $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ de rango 1 arriba de la diagonal.)

43. Todas las **matrices de Pascal** tienen determinante 1. Si 1 se resta del elemento n, n, ¿por qué el determinante se vuelve cero? (Use la regla 3, o cofactores).

$$\det \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & \mathbf{20} \end{bmatrix} = 1 \text{ conocido} \qquad \det \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 6 & 10 \\ 1 & 4 & 10 & \mathbf{19} \end{bmatrix} = \mathbf{0} \text{ (explique).}$$

4.4 APLICACIONES DE LOS DETERMINANTES

En esta sección se desarrollan cuatro aplicaciones fundamentales: la *inversa de A*, la *solución de $Ax = b$*, el *volumen de cajas* y los *pivotes*. Estas aplicaciones se encuentran entre los cálculos clave en álgebra lineal (realizados por eliminación). Los determinantes proporcionan fórmulas para las respuestas.

1. **Cálculo de A^{-1}.** El caso de 2 por 2 ilustra cómo los cofactores van en A^{-1}:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{\det A}\begin{bmatrix} C_{11} & C_{21} \\ C_{12} & C_{22} \end{bmatrix}.$$

Se está dividiendo entre el determinante, y A es invertible exactamente cuando $\det A$ es diferente de cero. El número $C_{11} = d$ es el cofactor de a. El número $C_{12} = -c$ es el cofactor de b (observe el signo menos). ¡Ese número C_{12} va en el renglón 2, columna 1!

El renglón a, b multiplicado por la columna C_{11}, C_{12} produce $ad - bc$. Este es el desarrollo por cofactores de $\det A$. Esta es la pista que se requiere: **A^{-1} divide los cofactores entre det A.**

La matriz de cofactores C está traspuesta $\quad A^{-1} = \dfrac{C^{\mathrm{T}}}{\det A}$ significa $(A^{-1})_{ij} = \dfrac{C_{ji}}{\det A}$. (1)

Nuestro objetivo es comprobar esta fórmula para A^{-1}. Es necesario ver por qué $AC^{\mathrm{T}} = (\det A)I$:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}\begin{bmatrix} C_{11} & \cdots & C_{n1} \\ \vdots & & \vdots \\ C_{1n} & \cdots & C_{nn} \end{bmatrix} = \begin{bmatrix} \det A & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \det A \end{bmatrix}. \tag{2}$$

Con los cofactores $C_{11}, \ldots, C_{1n}$ en la primera *columna* y no en el primer renglón, se multiplican a $a_{11}, \ldots, a_{1n}$ y proporcionan el elemento diagonal $\det A$. Cada renglón de A multiplica a sus cofactores (*el desarrollo por cofactores*) para obtener la misma respuesta $\det A$ sobre la diagonal.

La pregunta crítica es: *¿Por qué se obtienen ceros fuera de la diagonal*? Si se combinan los elementos a_{1j} del renglón 1 con los cofactores C_{2j} del renglón 2, ¿por qué el resultado es cero?

renglón 1 de A, renglón 2 de C $\quad a_{11}C_{21} + a_{12}C_{22} + \cdots + a_{1n}C_{2n} = 0$. (3)

La respuesta es: se está calculando el determinante de una nueva matriz B, con un nuevo renglón 2. El primer renglón de A se copia en el segundo renglón de B. Así, B tiene dos renglones iguales, y $\det B = 0$. La ecuación (3) es el desarrollo de $\det B$ a lo largo de su renglón 2, donde B tiene exactamente los mismos cofactores que A (porque el segundo renglón se elimina para encontrar esos cofactores). La extraordinaria multiplicación matricial (2) es correcta.

Esa multiplicación $AC^{\mathrm{T}} = (\det A)I$ proporciona de inmediato A^{-1}. Recuerde que el cofactor al borrar el renglón i y la columna j de A va en el *renglón j y en la columna i* de C^{T}. Al dividir entre el número $\det A$ (¡en caso de no ser cero!) se obtiene $A^{-1} = C^{\mathrm{T}}/\det A$.

Ejemplo 1 La inversa de una suma de matrices es una matriz en diferencias:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \text{ tiene } A^{-1} = \frac{C^{\mathrm{T}}}{\det A} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

El signo menos entra porque los cofactores siempre incluyen a $(-1)^{i+j}$.

2. **La solución de $Ax = b$.** La multiplicación $x = A^{-1}b$ es justo $C^{\mathrm{T}}b$ dividido entre $\det A$. Hay una forma conocida de escribir la respuesta $(x_1, \ldots, x_n)$:

4C ***Regla de Cramer:*** La j-ésima componente de $x = A^{-1}b$ es el cociente

$$x_j = \frac{\det B_j}{\det A}, \quad \text{donde} \quad B_j = \begin{bmatrix} a_{11} & a_{12} & b_1 & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & b_n & a_{nn} \end{bmatrix} \text{ tiene } b \text{ en la columna } j. \tag{4}$$

Demostración B_j se desarrolla por cofactores de su columna j-ésima (que es b). Como los cofactores ignoran esa columna, det B_j es exactamente la j-ésima componente en el producto C^Tb:

$$\det B_j = b_1C_{1j} + b_2C_{2j} + \cdots + b_nC_{nj}.$$

Al dividir lo anterior entre det A se obtiene x_j. Cada componente de x es un *cociente de dos determinantes*. Este hecho pudo haberse reconocido a partir de la eliminación gaussiana, pero no fue así. ■

Ejemplo 2 La solución de

$$\begin{aligned} x_1 + 3x_2 &= 0 \\ 2x_1 + 4x_2 &= 6 \end{aligned}$$

tiene 0 y 6 en la primera columna para x_1, y en la segunda columna para x_2:

$$x_1 = \frac{\begin{vmatrix} 0 & 3 \\ 6 & 4 \end{vmatrix}}{\begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix}} = \frac{-18}{-2} = 9, \qquad x_2 = \frac{\begin{vmatrix} 1 & 0 \\ 2 & 6 \end{vmatrix}}{\begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix}} = \frac{6}{-2} = -3.$$

Los denominadores siempre son det A. Para 1 000 ecuaciones, la regla de Cramer requiere 1 001 determinantes. Para mi consternación, encontré en un libro titulado *Mathematics for the Millions* que la regla de Cramer era realmente recomendada (y la eliminación se dejaba de lado):

> Para tratar con un conjunto que implica cuatro variables u, w, z, primero es necesario eliminar una de ellas en cada uno de los tres pares con la finalidad de obtener tres ecuaciones en tres variables y así proceder para el triple miembro izquierdo para obtener valores para dos de ellas. El lector que haga esto como ejercicio empezará a darse cuenta de cuán formidablemente laborioso se convierte el método de eliminación, donde es necesario tratar con más de tres variables. Esta consideración nos invita a explorar la posibilidad de un *método más rápido* . . .

¡El "método más rápido" es la regla de Cramer! Si el autor intentaba calcular 1 001 determinantes, ¡yo llamaría al libro *Mathematics for the Millionaire*!

3. El volumen de una caja. La relación entre el determinante y el volumen se vuelve más evidente cuando los *ángulos* son *rectos*: las aristas son perpendiculares, y la caja es rectangular. Así, el volumen es el producto de las longitudes de las aristas: *volumen* $= \ell_1\ell_2\cdots\ell_n$.

Se desea obtener el mismo $\ell_1\ell_2\cdots\ell_n$ a partir de det A, *cuando las aristas de esa caja estén en los renglones de* A. Con ángulos rectos, estos renglones son ortogonales y AA^T es diagonal:

Caja con ángulos rectos
Renglones ortogonales

$$AA^T = \begin{bmatrix} \text{renglón } 1 \\ \vdots \\ \text{renglón } n \end{bmatrix} \begin{bmatrix} \text{r} & & \text{r} \\ \text{e} & & \text{e} \\ \text{n} & & \text{n} \\ \text{g} & & \text{g} \\ \text{l} & \cdots & \text{l} \\ \text{ó} & & \text{ó} \\ \text{n} & & \text{n} \\ 1 & & n \end{bmatrix} = \begin{bmatrix} \ell_1^2 & & 0 \\ & \ddots & \\ 0 & & \ell_n^2 \end{bmatrix}.$$

Las ℓ_i son las longitudes de los renglones (las aristas), y los ceros fuera de la diagonal se deben a que los renglones son ortogonales. Usando las reglas del producto y de la trasposición,

Caso con ángulos rectos $\quad \ell_1^2\ell_2^2\cdots l_n^2 = \det(AA^T) = (\det A)(\det A^T) = (\det A)^2.$

La raíz cuadrada de esta ecuación indica que ***el determinante es igual al volumen.*** El *signo* de det A indica si las aristas constituyen un conjunto "derecho" de coordenadas, como en el sistema x-y-z de costumbre, o un sistema izquierdo, como y-x-z.

Si los ángulos no son de 90°, entonces el volumen no es el producto de las longitudes. En el plano (véase la figura 4.2), el "volumen" de un paralelogramo es igual a la base ℓ multiplicada por la altura h. El vector $b - p$ de longitud h es el segundo renglón $b = (a_{21}, a_{22})$, menos su proyección p sobre el primer renglón. La cuestión clave es ésta: Por la regla 5, det A permanece sin cambio cuando un múltiplo del renglón 1 se resta del renglón 2. *El paralelogramo puede cambiar a un rectángulo*, donde ya se ha demostrado que volumen = determinante.

En n dimensiones, se requiere más trabajo para hacer rectangular a cada caja, aunque la idea es la misma. El volumen y el determinante permanecen sin cambio si de cada renglón se resta su proyección sobre el espacio generado por los renglones precedentes, dejando un "vector longitud" perpendicular como pb. Este proceso de Gram-Schmidt produce renglones ortogonales, con volumen = determinante. Así, la misma desigualdad debe cumplirse para los renglones originales.

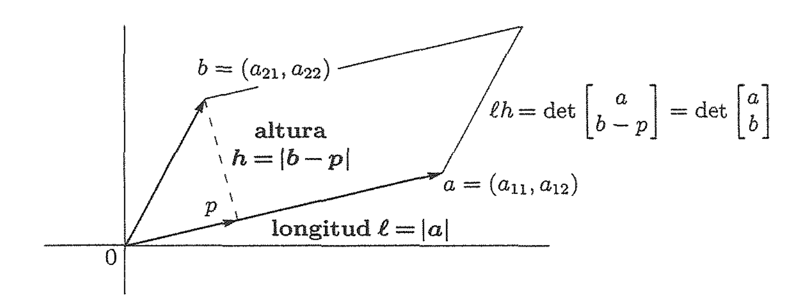


Figura 4.2 Volumen (área) del paralelogramo = ℓ veces h = |det A|.

Así se completa el vínculo entre los volúmenes y los determinantes, aunque merece la pena volver una vez más al caso más simple. Se sabe que

$$\det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1, \qquad \det \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} = 1.$$

Estos determinantes proporcionan los volúmenes, o las áreas, ya que se está trabajando en dos dimensiones, que se muestran en la figura 4.3. La base y la altura del paralelogramo miden uno; de modo que su área también es 1.

4. Una fórmula para los pivotes. Finalmente ya es posible saber cuándo es posible realizar la eliminación sin intercambios de renglones. La observación clave es que los k primeros pivotes están determinados completamente por la submatriz A_k en la esquina superior izquierda de A. *Los renglones y las columnas restantes de A no afectan esta*

Figura 4.3 Las áreas de un cuadrado unitario y de un paralelogramo unitario son iguales a 1.

esquina del problema:

La eliminación sobre A incluye la eliminación sobre A_2

$$A = \begin{bmatrix} a & b & e \\ c & d & f \\ g & h & i \end{bmatrix} \rightarrow \begin{bmatrix} a & b & e \\ \mathbf{0} & (ad - bc)/a & (af - ec)/a \\ g & h & i \end{bmatrix}.$$

Ciertamente, el primer pivote sólo depende del primer renglón y de la primera columna. El segundo pivote $(ad - bc)/a$, sólo depende de la submatriz esquinada A_2 de 2 por 2. El resto de A no entra sino hasta el tercer pivote. En realidad, lo que determina la esquina superior izquierda de L no sólo son los pivotes, sino todas las esquinas superiores izquierdas de L, D, y U:

$$A = LDU = \begin{bmatrix} 1 & & \\ c/a & 1 & \\ * & * & 1 \end{bmatrix} \begin{bmatrix} a & & \\ & (ad-bc)/a & \\ & & * \end{bmatrix} \begin{bmatrix} 1 & b/a & * \\ & 1 & * \\ & & 1 \end{bmatrix}.$$

Lo que se ve en los dos primeros renglones y columnas es exactamente la factorización de la submatriz esquinada A_2. Esta es una regla general si no hay intercambios de renglones:

> **4D** Si A se factoriza en LDU, las esquinas superiores izquierdas cumplen $A_k = L_k D_k U_k$. Para toda k, la submatriz A_k experimenta eliminación gaussiana de suyo.

La demostración es ver que esta esquina puede establecerse primero, incluso antes de considerar otras eliminaciones. O bien, se aplican las reglas para la ***multiplicación por bloques***:

$$LDU = \begin{bmatrix} L_k & 0 \\ B & C \end{bmatrix} \begin{bmatrix} D_k & 0 \\ 0 & E \end{bmatrix} \begin{bmatrix} U_k & F \\ 0 & G \end{bmatrix} = \begin{bmatrix} L_k D_k U_k & L_k D_k F \\ B D_k U_k & B D_k F + CEG \end{bmatrix}.$$

Al comparar la última matriz con A, la esquina $L_k D_k U_k$ coincide con A_k. Así,

$$\det A_k = \det L_k \det D_k \det U_k = \det D_k = d_1 d_2 \cdots d_k.$$

El producto de los k primeros pivotes es el determinante de A_k. Ésta es la misma regla que ya se conoce para toda matriz. Debido a que el determinante de A_{k-1} está dado por $d_1 d_2 \cdots d_{k-1}$, cada pivote d_k puede aislarse como un ***cociente de determinantes***:

Fórmula para los pivotes
$$\frac{\det A_k}{\det A_{k-1}} = \frac{d_1 d_2 \cdots d_k}{d_1 d_2 \cdots d_{k-1}} = d_k. \qquad (5)$$

En el ejemplo anterior, el segundo pivote era exactamente este cociente $(ad - bc)/a$. Es el determinante de A_2 dividido entre el determinante de A_1. (Por convencionalismo, $\det A_0 = 1$, de modo que el primer pivote es $a/1 = a$).

Al multiplicar entre sí a todos los pivotes individuales, se recupera

$$d_1 d_2 \cdots d_n = \frac{\det A_1}{\det A_0} \frac{\det A_2}{\det A_1} \cdots \frac{\det A_n}{\det A_{n-1}} = \frac{\det A_n}{\det A_0} = \det A.$$

Con base en la ecuación (5) es posible leer finalmente la respuesta de nuestra pregunta original: ***Todos los elementos pivote son diferentes de cero siempre que todos los números* $\det A_k$ *sean diferentes de cero***:

> **4E** La eliminación puede completarse sin intercambios de renglones (de modo que $P = I$ y $A = LU$), si y sólo si todas las submatrices principales $A_1, A_2, \ldots, A_n$ son no singulares.

Eso es lo concerniente a los determinantes, excepto por una observación opcional sobre la propiedad 2: el cambio de signo sobre los intercambios de renglones. ***El determinante de una matriz permutación P*** es el único punto cuestionable en la gran fórmula. Independientemente de los intercambios particulares en los renglones que vinculan P con I, el número de intercambios ¿siempre es par o impar? De ser así, su determinante está bien definido por la regla 2, como $+1$ o -1.

Empezando con (3, 2, 1), con un simple intercambio de 3 y 1 podría llegarse al orden natural (1, 2, 3). Así sería también con un intercambio de 3 y 2, luego de 3 y 1, y después de 2 y 1. En ambas secuencias, el número de intercambios es impar. La afirmación es que *un número par de intercambios jamás puede producir el orden natural, empezando con* (3, 2, 1).

A continuación se proporciona una demostración. Considere cada par de números en la permutación, y sea N la cantidad de pares en que primero aparece el número mayor. Ciertamente, $N = 0$ para el orden natural (1, 2, 3). El orden (3, 2, 1) tiene $N = 3$ porque todos los pares (3, 2), (3, 1), y (2, 1) son erróneos. Se demostrará que *todo intercambio modifica a N por un número impar*. Luego, para llegar a $N = 0$ (el orden natural) se requiere un número de intercambios que tenga la misma característica par o impar que N.

Cuando se intercambian vecinos, N cambia por $+1$ o -1. *Cualquier intercambio puede lograrse mediante un número impar de intercambios de vecinos*. Esto completa la demostración; un número impar de números impares es impar. Para intercambiar los elementos primero y cuarto a continuación, que son 2 y 3, se requieren cinco intercambios (un número impar) de vecinos:

$$(\mathbf{2}, 1, 4, \mathbf{3}) \rightarrow (1, \mathbf{2}, 4, \mathbf{3}) \rightarrow (1, 4, \mathbf{2}, \mathbf{3}) \rightarrow (1, 4, \mathbf{3}, \mathbf{2}) \rightarrow (1, \mathbf{3}, 4, \mathbf{2}) \rightarrow (\mathbf{3}, 1, 4, \mathbf{2}).$$

Se requieren $\ell - k$ intercambios de vecinos para mover el elemento que está en el lugar k al lugar ℓ. Luego, $\ell - k - 1$ intercambios mueven al elemento que originalmente estaba en el sitio ℓ (y que ahora se encuentra en el lugar $\ell - 1$) de regreso al lugar k. Debido a que $(\ell - k) + (\ell - k - 1)$ es impar, se ha completado la demostración. El determinante no sólo tiene todas las propiedades antes encontradas, incluso existe.

Conjunto de problemas 4.4

1. Encuentre el determinante y todos los nueve cofactores C_{ij} de la siguiente matriz triangular:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix}.$$

Escriba C^T, y compruebe que $AC^T = (\det A)I$. ¿Qué es A^{-1}?

2. Use la matriz de cofactores C para invertir las siguientes matrices simétricas:

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

3. Encuentre x, y, y z aplicando la regla de Cramer en la ecuación (4):

$$\begin{aligned} ax + by &= 1 \\ cx + dy &= 0 \end{aligned} \quad \text{y} \quad \begin{aligned} x + 4y - z &= 1 \\ x + y + z &= 0 \\ 2x \qquad + 3z &= 0. \end{aligned}$$

4. *a*) Encuentre el determinante cuando un vector x sustituye a la columna j de la identidad (considere $x_j = 0$ como un caso por separado):

$$\text{si} \quad M = \begin{bmatrix} 1 & & x_1 & & \\ & 1 & \cdot & & \\ & & x_j & & \\ & & \cdot & 1 & \\ & & x_n & & 1 \end{bmatrix} \text{entonces} \quad \det M = ___.$$

b) Si $Ax = b$, demuestre que AM es la matriz B_j en la ecuación (4), con b en la columna j.

c) Deduzca la *regla de Cramer*, tomando determinantes en $AM = B_j$.

5. *a*) Dibuje el triángulo con vértices $A = (2, 2)$, $B = (-1, 3)$, y $C = (0, 0)$. Al considerarlo como la mitad de un paralelogramo, explique por qué su área es igual a

$$\text{área } (ABC) = \frac{1}{2} \det \begin{bmatrix} 2 & 2 \\ -1 & 3 \end{bmatrix}.$$

b) Mueva el tercer vértice a $C = (1, -4)$ y justifique la fórmula

$$\text{área } (ABC) = \frac{1}{2} \det \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix} = \frac{1}{2} \det \begin{bmatrix} 2 & 2 & 1 \\ -1 & 3 & 1 \\ 1 & -4 & 1 \end{bmatrix}.$$

Sugerencia: Al restar el último renglón de cada uno de los demás se obtiene

$$\det \begin{bmatrix} 2 & 2 & 1 \\ -1 & 3 & 1 \\ 1 & -4 & 1 \end{bmatrix} = \det \begin{bmatrix} 1 & 6 & 0 \\ -2 & 7 & 0 \\ 1 & -4 & 1 \end{bmatrix} = \det \begin{bmatrix} 1 & 6 \\ -2 & 7 \end{bmatrix}.$$

Trace $A' = (1, 6)$, $B' = (-2, 7)$, $C' = (0, 0)$ y su relación con A, B, C.

6. Explique en términos de volúmenes por qué $\det 3A = 3^n \det A$ para cualquier matriz A de n por n.

7. Pronostique, y confirme por eliminación, los elementos pivotes de

$$A = \begin{bmatrix} 2 & 1 & 2 \\ 4 & 5 & 0 \\ 2 & 7 & 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 2 & 1 & 2 \\ 4 & 5 & 3 \\ 2 & 7 & 0 \end{bmatrix}.$$

8. Encuentre todas las permutaciones impares de los números $\{1, 2, 3, 4\}$. Provienen de un número impar de intercambios y conduce a $\det P = -1$.

9. Suponga que la permutación P lleva $(1, 2, 3, 4, 5)$ a $(5, 4, 1, 2, 3)$.

a) ¿Qué hace P^2 a $(1, 2, 3, 4, 5)$?

b) ¿Qué hace P^{-1} a $(1, 2, 3, 4, 5)$?

10. Si P es una permutación impar, explique por qué P^2 es par pero P^{-1} es impar.

11. Demuestre que si se mantiene la multiplicación de A por la misma matriz permutación P, entonces el primer renglón termina por regresar a su sitio original.

12. Si A es una matriz de 5 por 5 con todos $|a_{ij}| \leq 1$, entonces $\det A \leq$ ____. Los volúmenes, la gran fórmula o los pivotes deben proporcionar alguna cota superior para el determinante.

Los problemas 13 al 17 son sobre la regla de Cramer para $x = A^{-1}b$.

13. Resuelva las siguientes ecuaciones lineales, aplicando la regla de Cramer $x_j = \det B_j / \det A$:

$$a) \quad \begin{aligned} 2x_1 + 5x_2 &= 1 \\ x_1 + 4x_2 &= 2. \end{aligned} \qquad b) \quad \begin{aligned} 2x_1 + x_2 \qquad &= 1 \\ x_1 + 2x_2 + x_3 &= 0 \\ x_2 + 2x_3 &= 0. \end{aligned}$$

14. Use la regla de Cramer para despejar (solamente) y. Sea D el determinante de 3 por 3:

$$a) \quad \begin{aligned} ax + by &= 1 \\ cx + dy &= 0. \end{aligned} \qquad b) \quad \begin{aligned} ax + by + cz &= 1 \\ dx + ey + fz &= 0 \\ gx + hy + iz &= 0. \end{aligned}$$

15. La regla de Cramer falla cuando $\det A = 0$. El ejemplo a) no tiene solución, mientras *b*) tiene una infinidad. ¿Cuáles son los cocientes $x_j = \det B_j / \det A$?

$$a) \quad \begin{aligned} 2x_1 + 3x_2 &= 1 \\ 4x_1 + 6x_2 &= 1. \end{aligned} \text{ (rectas paralelas)} \qquad b) \quad \begin{aligned} 2x_1 + 3x_2 &= 1 \\ 4x_1 + 6x_2 &= 2. \end{aligned} \text{ (la misma recta)}$$

16. *Demostración rápida de la regla de Cramer*. El determinante es una función lineal de la columna 1. Es cero si dos columnas son iguales. Cuando $\boldsymbol{b} = A\boldsymbol{x} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3$ va en la columna 1 para producir B_1, el determinante es

$$\begin{vmatrix} \boldsymbol{b} & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{vmatrix} = \begin{vmatrix} x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{vmatrix} = x_1 \begin{vmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{vmatrix} = x_1 \det A.$$

a) ¿Qué fórmula para x_1 proviene de miembro izquierdo = miembro derecho?

b) ¿Qué pasos llevan a la ecuación de en medio?

17. Si el miembro derecho b es la *última columna* de A, resuelva el sistema $Ax = b$ de 3 por 3. Explique cómo cada determinante en la regla de Cramer conduce a su solución x.

Los problemas 18 a 26 son sobre $A^{-1} = C^{\mathrm{T}}/\det A$. Recuerde trasponer C.

18. Encuentre A^{-1} a partir de la fórmula de cofactores $C^{\mathrm{T}}/\det A$. Use simetría en el inciso *b*):

$$a) \quad A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 0 \\ 0 & 4 & 1 \end{bmatrix}. \qquad b) \quad A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}.$$

19. Si todos los cofactores son cero, ¿cómo se sabe que A no tiene inversa? Si ninguno de los cofactores es cero, ¿es seguro que A es invertible?

20. Encuentre los cofactores de A, y multiplique AC^{T} para encontrar $\det A$:

$$A = \begin{bmatrix} 1 & 1 & 4 \\ 1 & 2 & 2 \\ 1 & 2 & 5 \end{bmatrix}, \qquad C = \begin{bmatrix} 6 & -3 & 0 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}, \qquad \text{y} \qquad AC^{\mathrm{T}} = _____.$$

Si el elemento de la esquina se cambia de 4 a 100, ¿por qué $\det A$ permanece sin cambio?

21. Suponga que $\det A = 1$ y que se conocen todos los cofactores. ¿Cómo puede encontrarse A?

22. A partir de la fórmula $AC^T = (\det A)I$, demuestre que $\det C = (\det A)^{n-1}$.

23. (Sólo para profesores) Si se conocen todos los 16 cofactores de una matriz invertible A de 4 por 4, ¿cómo encontrar A?

24. Si todos los elementos de A son enteros, y $\det A = 1$ o -1, demuestre que todos los elementos de A^{-1} son enteros. Proporcione un ejemplo de 2 por 2.

25. L es triangular inferior y S es simétrica. Suponga que son invertibles:

$$L = \begin{bmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{bmatrix} \qquad S = \begin{bmatrix} a & b & d \\ b & c & e \\ d & e & f \end{bmatrix}.$$

a) ¿Cuáles son los tres cofactores de L que son cero? Entonces L^{-1} es triangular inferior.

b) ¿Cuáles son los tres pares de cofactores de S que son iguales? Entonces S^{-1} es simétrica.

26. Para $n = 5$, la matriz C contiene ____ cofactores y cada cofactor de 4 por 4 contiene ____ términos y cada término requiere ____ multiplicaciones. Compare con $5^3 = 125$ para el cálculo de Gauss-Jordan de A^{-1}.

Los problemas 27 a 36 son sobre área y volumen por determinantes.

27. *a)* Encuentre el área del paralelogramo con aristas $v = (3, 2)$ y $w = (1, 4)$.
b) Encuentre el área del triángulo con lados v, w, y $v + w$. Dibújelo.
c) Encuentre el área del triángulo con lados v, w, y $w - v$. Dibújelo.

28. Las aristas de una caja van de $(0, 0, 0)$ a $(3, 1, 1)$, $(1, 3, 1)$, y $(1, 1, 3)$. Encuentre su volumen y también el área de cada cara del paralelogramo.

29. *a)* Los vértices de un triángulo son $(2, 1)$, $(3, 4)$, y $(0, 5)$. ¿Cuál es el área?
b) Un nuevo vértice en $(-1, 0)$ lo hace de lobular (cuatro lados). Encuentre el área.

30. El paralelogramo con lados $(2, 1)$ y $(2, 3)$ tiene la misma área que el paralelogramo con lados $(2, 2)$ y $(1, 3)$. Encuentre estas áreas a partir de determinantes de 2 por 2, y justifique por qué deben ser iguales. (No puedo verlo en una ilustración. Por favor escríbame si usted puede verlo.)

31. La matriz H de Hadamard tiene renglones ortogonales. ¡La caja es un hipercubo!

$$¿\text{Qué es } \det H = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{vmatrix} = \text{volumen de un hipercubo en } \mathbf{R}^4?$$

32. Si las longitudes de las columnas de una matriz de 4 por 4 son L_1, L_2, L_3, L_4, ¿cuál es el máximo valor posible para el determinante (con base en el volumen)? Si todos los elementos son 1 o -1, ¿cuáles son esas longitudes y el determinante máximo?

33. Demuestre con una figura cómo un rectángulo con área x_1y_2 menos un rectángulo con área x_2y_1 produce el área $x_1y_2 - x_2y_1$ de un paralelogramo.

34. Cuando los vectores arista $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ son perpendiculares, el volumen de la caja es $\|\boldsymbol{a}\|$ por $\|\boldsymbol{b}\|$ por $\|\boldsymbol{c}\|$. La matriz A^TA es ____. Encuentre $\det A^TA$ y $\det A$.

35. ¿Cuántos vértices tiene un cubo n-dimensional? ¿Cuántas aristas? ¿Cuántas caras $n - 1$ dimensionales? El n-cubo cuyas aristas son los renglones de $2I$ tiene volumen ____. Una computadora hipercúbica tiene procesadores en paralelo en los vértices, con conexiones a lo largo de las aristas.

36. El área de un triángulo con vértices $(0, 0)$, $(1, 0)$, $(0, 1)$ es $\frac{1}{2}$. El volumen de la pirámide con cuatro vértices $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ es ____. ¿Cuál es el volumen de la pirámide en $\mathbf{R}^4$ con cinco vértices en $(0, 0, 0, 0)$ y los renglones de I?

Los problemas 37 a 40 son sobre áreas dA y volúmenes dV, en cálculo.

37. Las coordenadas polares satisfacen $x = r\cos\theta$ y $y = r \operatorname{sen}\theta$. El área polar $J\,dr\,d\theta$ incluye a J:

$$J = \begin{vmatrix} \partial x/\partial r & \partial x/\partial \theta \\ \partial y/\partial r & \partial y/\partial \theta \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\operatorname{sen}\theta \\ \operatorname{sen}\theta & r\cos\theta \end{vmatrix}.$$

Las dos columnas son ortogonales. Sus longitudes son ____. Así, $J =$ ____.

38. Las coordenadas esféricas ρ, ϕ, θ proporcionan $x = \rho \operatorname{sen}\phi \cos\theta$, $y = \rho \operatorname{sen}\phi \operatorname{sen}\theta$, $z = \rho\cos\phi$. Encuentre la matriz jacobiana de 9 derivadas parciales: $\partial x/\partial\rho, \partial x/\partial\phi, \partial x/\partial\theta$ están en el renglón 1. Simplifique su determinante a $J = \rho^2 \operatorname{sen}\phi$. Luego, $dV = \rho^2 \operatorname{sen}\phi\, d\rho\, d\phi\, d\theta$.

39. La matriz que relaciona r, θ con x, y y está en el problema 37. Invierta esa matriz:

$$J^{-1} = \begin{vmatrix} \partial r/\partial x & \partial r/\partial y \\ \partial \theta/\partial x & \partial \theta/\partial y \end{vmatrix} = \begin{vmatrix} \cos\theta & ? \\ ? & ? \end{vmatrix} = ?$$

Es sorprendente que $\partial r/\partial x = \partial x/\partial r$. El producto $JJ^{-1} = I$ proporciona la regla de la cadena

$$\frac{\partial x}{\partial x} = \frac{\partial x}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial x}{\partial \theta}\frac{\partial \theta}{\partial x} = 1.$$

40. El área del triángulo con vértices $(0, 0)$, $(6, 0)$, y $(1, 4)$ es ____. Cuando se hace rotar un ángulo $\theta = 60°$, el área es ____. La matriz rotación tiene

$$\text{determinante } = \begin{vmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & ? \\ ? & ? \end{vmatrix} = ?$$

41. Sean $P = (1, 0, -1)$, $Q = (1, 1, 1)$, y $R = (2, 2, 1)$. Escoja S de modo que $PQRS$ sea un paralelogramo, y calcule su área. Escoja T, U, V de modo que $OPQRSTUV$ sea una caja inclinada, y calcule su volumen.

42. Suponga que (x, y, z), $(1, 1, 0)$, y $(1, 2, 1)$ están en un plano que pasa por el origen. ¿Cuál determinante es cero? ¿Qué ecuación proporciona esto para el plano?

43. Suponga que (x, y, z) es una combinación lineal de $(2, 3, 1)$ y $(1, 2, 3)$. ¿Cuál determinante es cero? ¿Qué ecuación proporciona esto para el plano de todas las combinaciones?

44. Si $Ax = (1, 0, \ldots, 0)$, demuestre que la regla de Cramer proporciona $x =$ *primera columna de* A^{-1}.

45. (VISA a AVIS) Esto requiere un número impar de intercambios (IVSA, AVSI, AVIS). Cuente los pares de letras en VISA y AVIS que están invertidas en orden alfabético. La diferencia debe ser impar.

Capítulo 4 Ejercicios de repaso

4.1 Encuentre los determinantes de

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \\ 1 & 1 & 3 & 1 \\ 1 & 4 & 1 & 1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 2 & -1 & 0 & -1 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ -1 & 0 & -1 & 2 \end{bmatrix}.$$

4.2 Si $B = M^{-1}AM$, ¿por qué es cierto que $\det B = \det A$? También demuestre que $\det A^{-1}B = 1$.

4.3 Empezando con A, multiplique su primer renglón por 3 para obtener B, y reste el primer renglón de B del segundo para obtener C. ¿Cómo está relacionado $\det C$ con $\det A$?

4.4 Resuelva $3u + 2v = 7$, $4u + 3v = 11$, aplicando la regla de Cramer.

4.5 Si todos los elementos de A y A^{-1} son enteros, ¿cómo sabe que ambos determinantes son 1 o -1? *Sugerencia*: ¿A qué es igual $\det A$ multiplicado por $\det A^{-1}$?

4.6 Encuentre todos los cofactores, y la inversa o el espacio nulo, de

$$\begin{bmatrix} 3 & 5 \\ 6 & 9 \end{bmatrix}, \quad \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}, \quad \text{y} \quad \begin{bmatrix} a & b \\ a & b \end{bmatrix}.$$

4.7 ¿Cuál es el volumen del paralelepípedo que tiene cuatro de sus vértices en $(0, 0, 0)$, $(-1, 2, 2)$, $(2, -1, 2)$, y $(2, 2, 1)$? ¿Dónde están los otros cuatro vértices?

4.8 ¿Cuántos términos hay en el desarrollo de un determinante de 5 por 5, y cuántos de éstos es seguro que son cero si $a_{21} = 0$?

4.9 Si P_1 es una matriz permutación par y P_2 es impar, a partir de $P_1 + P_2 = P_1(P_1^{\mathrm{T}} + P_2^{\mathrm{T}})P_2$ deduzca que $\det(P_1 + P_2) = 0$.

4.10 Si $\det A > 0$, demuestre que A puede conectarse a I mediante una cadena continua de matrices $A(t)$, todas con determinantes positivos. (La ruta directa $A(t) = A + t(I - A)$ va de $A(0) = A$ a $A(1) = I$, pero entretanto $A(t)$ puede ser singular. El problema no es tan sencillo, y el autor agradecerá las soluciones que le sean enviadas.)

4.11 Explique por qué el punto (x, y) está sobre la recta que pasa por $(2, 8)$ y $(4, 7)$, si

$$\det \begin{bmatrix} x & y & 1 \\ 2 & 8 & 1 \\ 4 & 7 & 1 \end{bmatrix} = 0, \quad \text{o bien} \quad x + 2y - 18 = 0.$$

4.12 En analogía con el ejercicio previo, ¿cuál es la ecuación para que (x, y, z) esté sobre el plano que pasa por $(2, 0, 0)$, $(0, 2, 0)$ y $(0, 0, 4)$? Implica un determinante de 4 por 4.

4.13 Si los puntos (x, y, z), $(2, 1, 0)$, y $(1, 1, 1)$ están sobre un plano que pasa por el origen, ¿qué determinante es cero? Los vectores $(1, 0, -1)$, $(2, 1, 0)$, $(1, 1, 1)$, ¿son independientes?

4.14 Si todo renglón de A tiene un solo $+1$, o un solo -1, o uno de cada uno (y en caso contrario es cero), demuestre que $\det A = 1$ o -1 o 0.

4.15 Si $C = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ y $D = \begin{bmatrix} u & v \\ w & z \end{bmatrix}$, entonces $CD = -DC$ produce 4 ecuaciones $Ax = 0$:

$$CD + DC = 0 \quad \text{es} \quad \begin{bmatrix} 2a & c & b & 0 \\ b & a+d & 0 & b \\ c & 0 & a+d & c \\ 0 & c & b & 2d \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

a) Demuestre que $\det A = 0$ si $a + d = 0$. Resuelva para u, v, w, z, los elementos de D.

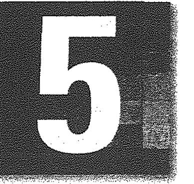
b) Demuestre que $\det A = 0$ si $ad = bc$ (de modo que C es singular).

En todos los demás casos, $CD = -DC$ sólo es posible con D = matriz cero.

4.16 El desplazamiento circular permuta $(1, 2, \ldots, n)$ en $(2, 3, \ldots, 1)$. ¿Cuál es la matriz permutación correspondiente P, y (dependiendo de n), cuál es su determinante?

4.17 Encuentre el determinante de A = eye(5) + ones(5) y si es posible, el de eye(n) + ones(n).

Capítulo



Valores característicos y vectores característicos

5.1 INTRODUCCIÓN

Este capítulo inicia la "segunda parte" del álgebra lineal. La primera mitad concernía a $Ax = b$. El nuevo problema, $Ax = \lambda x$, se resolverá simplificando una matriz, haciéndola diagonal de ser posible. *El paso fundamental es ya no restar un múltiplo de un renglón de otro*. La eliminación modifica los valores característicos, lo cual no es deseable.

Los determinantes proporcionan una transición de $Ax = b$ a $Ax = \lambda x$. En ambos casos el determinante lleva a una "solución formal": para la regla de Cramer para $x = A^{-1}b$, y para el polinomio det $(A - \lambda I.)$ cuyas raíces son los valores característicos. (Ahora todas las matrices son cuadradas; los valores característicos de una matriz rectangular no tienen más sentido que su determinante.) El determinante puede usarse realmente si $n = 2$ o 3. Para n grande, el cálculo de λ es más difícil que resolver $Ax = b$.

El primer paso es comprender la manera en que los valores característicos pueden ser de utilidad. Una de las aplicaciones de éstos es en las ecuaciones diferenciales. ¡Aquí se supone que el lector no es un experto en ecuaciones diferenciales! Si el lector puede diferenciar x^n, sen x, y e^x, sabe lo suficiente. Como ejemplo específico, considere el par de ecuaciones acopladas

$$\frac{dv}{dt} = 4v - 5w, \quad v = 8 \quad \text{en} \quad t = 0,$$
$$\frac{dw}{dt} = 2v - 3w, \quad w = 5 \quad \text{en} \quad t = 0. \tag{1}$$

Este es un ***problema con valor inicial***. La incógnita se especifica en el instante $t = 0$, mediante los valores iniciales proporcionados 8 y 5. El problema consiste en encontrar $v(t)$ y $w(t)$ para instantes posteriores $t > 0$.

Resulta fácil escribir el sistema en forma matricial. Sea $u(t)$ el vector que se desconoce, con valor inicial $u(0)$. La matriz de coeficientes es A:

Vector desconocido $\quad u(t) = \begin{bmatrix} v(t) \\ w(t) \end{bmatrix}, \quad u(0) = \begin{bmatrix} 8 \\ 5 \end{bmatrix}, \quad A = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix}.$

Las dos ecuaciones acopladas se convierten en la ecuación vectorial que se busca:

Forma matricial $$\frac{du}{dt} = Au \quad \text{con} \quad u = u(0) \text{ en } t = 0. \tag{2}$$

Este es el planteamiento básico del problema. Observe que se trata de una ecuación de primer orden; no aparecen derivadas de orden superior, y *es lineal* en las incógnitas. También tiene *coeficientes constantes*; la matriz A es independiente del tiempo.

¿Cómo se encuentra $u(t)$? Si en vez de una sola incógnita hubieran dos, sería fácil contestar esta pregunta. En lugar de una ecuación escalar se tendría una ecuación vectorial:

Ecuación simple $$\frac{du}{dt} = au \quad \text{con} \quad u = u(0) \text{ en } t = 0. \qquad (3)$$

Basta conocer la solución de esta ecuación:

Exponencial pura $$u(t) = e^{at}u(0). \qquad (4)$$

En el instante inicial $t = 0$, u es igual a $u(0)$ porque $e^0 = 1$. La derivada de e^{at} tiene el factor requerido a, de modo que $du/dt = au$. De esta forma se satisfacen tanto la condición inicial como la ecuación.

Observe el comportamiento de u para grandes instantes. La ecuación es inestable si $a > 0$, neutralmente estable si $a = 0$, o estable si $a < 0$; el factor e^{at} tiende al infinito, permanece acotado o tiende a cero. Si a fuese un número complejo, $a = \alpha + i\beta$, entonces las mismas pruebas podrían aplicarse a la parte real α. La parte compleja produce oscilaciones $e^{i\beta t} = \cos \beta t + i \operatorname{sen} \beta t$. La disminución o el crecimiento están regidos por el factor $e^{\alpha t}$.

Tanto para una simple ecuación. Se asumirá un método directo a sistemas, y se buscarán soluciones con la *misma dependencia exponencial sobre t* justo para encontrar en el caso escalar:

$$\begin{aligned} v(t) &= e^{\lambda t}y \\ w(t) &= e^{\lambda t}z \end{aligned} \qquad (5)$$

o, en notación vectorial,

$$u(t) = e^{\lambda t}x. \qquad (6)$$

Esta es toda la clave para las ecuaciones diferenciales $du/dt = Au$: ***buscar soluciones exponenciales puras***. Al sustituir $v = e^{\lambda t}y$ y $w = e^{\lambda t}z$ en la ecuación, se encuentra

$$\begin{aligned} \lambda e^{\lambda t}y &= 4e^{\lambda t}y - 5e^{\lambda t}z \\ \lambda e^{\lambda t}z &= 2e^{\lambda t}y - 3e^{\lambda t}z. \end{aligned}$$

El factor $e^{\lambda t}$ es común a cada término, por lo que puede eliminarse. Esta cancelación es la razón por la cual para ambas incógnitas se supone el mismo exponente λ; así se queda con

Problema de valor característico $$\begin{aligned} 4y - 5z &= \lambda y \\ 2y - 3z &= \lambda z. \end{aligned} \qquad (7)$$

Esta es la ecuación con valor característico. En forma matricial, se trata de $Ax = \lambda x$. Puede verla de nuevo si se utiliza $u = e^{\lambda t}x$: un número $e^{\lambda t}$ que crece o disminuye multiplicado por un vector fijo x. ***Al sustituir en du/dt = Au se obtiene*** $\lambda e^{\lambda t}x = Ae^{\lambda t}x$. ***Al cancelar*** $e^{\lambda t}$ ***se obtiene***

Ecuación de valor característico $$Ax = \lambda x. \qquad (8)$$

Ahora se cuenta con la ecuación fundamental de este capítulo. Implica dos incógnitas, λ y x. Se trata de un problema algebraico, ¡por lo que es posible olvidarse de las ecuaciones diferenciales! El número λ (lambda) es un ***valor característico*** de la matriz A, y el vector x es el ***vector característico*** asociado. Nuestro objetivo es encontrar los valores característicos y los vectores característicos, las λs y las xs, y utilizarlas.

Las soluciones de $Ax = \lambda x$

Observe que $Ax = \lambda x$ es una ecuación no lineal; λ multiplica a x. Si fuese posible encontrar λ, entonces la ecuación para x sería lineal. De hecho, en lugar de λx podría escribirse λIx, y pasar este término al miembro izquierdo:

$$(A - \lambda I)x = 0. \qquad (9)$$

La matriz identidad preserva matrices y preserva rectos a los vectores; la ecuación $(A - \lambda)x = 0$ es más corta, pero está mezclada. Esta es la clave del problema:

> ***El vector x está en el espacio nulo de $A - \lambda I$.***
> ***El número λ se escoge de modo que $A - \lambda I$ tenga un espacio nulo.***

Por supuesto, toda matriz tiene un espacio nulo. Resultaría ridículo sugerir lo contrario, pero el lector puede darse cuenta de este asunto. Se busca un vector característico *x distinto de cero*. El vector $x = 0$ siempre satisface $Ax = \lambda x$, aunque resulta inútil para resolver ecuaciones diferenciales. El objetivo es construir $u(t)$ sin exponenciales $e^{\lambda t}x$, y *se está interesado sólo en aquellos valores particulares λ para los que exista un vector característico x distinto de cero*. Para que sea útil, el espacio nulo de $A - \lambda I$ debe contener vectores diferentes de cero. En breve, $A - \lambda I$ ***debe ser singular***.

Para el efecto, el determinante proporciona una prueba concluyente.

5A El número λ es un valor característico de A si y sólo si $A - \lambda I$ es singular:

$$\det(A - \lambda I) = 0. \qquad (10)$$

Esta es la ecuación característica. Cada λ está asociada con vectores característicos x:

$$(A - \lambda I)x = 0 \quad \text{o bien,} \quad Ax = \lambda x. \qquad (11)$$

En nuestro ejemplo, A se sustituye por λI para hacerla singular:

Restar λI $$A - \lambda I = \begin{bmatrix} 4 - \lambda & -5 \\ 2 & -3 - \lambda \end{bmatrix}.$$

Observe que λ sólo se resta de la diagonal principal (ya que multiplica a I).

Determinante $$|A - \lambda I| = (4 - \lambda)(-3 - \lambda) + 10 \quad \text{o} \quad \lambda^2 - \lambda - 2.$$

Este es el *polinomio característico*. Sus raíces, donde el determinante es cero, son los valores característicos. Éstos provienen de la fórmula general para encontrar las raíces de una cuadrática, o de la factorización de $\lambda^2 - \lambda - 2 = (\lambda + 1)(\lambda - 2)$. Esto es cero si $\lambda = -1$

o $\lambda = 2$, lo cual se confirma con la fórmula general:

$$\textbf{Valores característicos} \qquad \lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{1 \pm \sqrt{9}}{2} = -1 \ \text{y} \ 2.$$

Hay dos valores característicos, porque una cuadrática tiene dos raíces. Toda matriz de 2 por 2 $A - \lambda I$ tiene a λ^2 (y ninguna otra potencia superior de λ) en su determinante.

Los valores $\lambda = -1$ y $\lambda = 2$ llevan a la solución de $Ax = \lambda x$ o $(A - \lambda I)x = 0$. Una matriz con determinante igual a cero es singular, por lo que en su espacio nulo debe haber vectores x diferentes de cero. De hecho, el espacio nulo contiene toda una *recta* de vectores característicos; ¡es un subespacio!

$$\lambda_1 = -1: \qquad (A - \lambda_1 I)x = \begin{bmatrix} 5 & -5 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

La solución (el primer vector característico) es cualquier múltiplo diferente de cero de x_1:

$$\textbf{Vector característico para } \lambda_1 \qquad x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

El cálculo de λ_2 se realiza por separado:

$$\lambda_2 = 2: \qquad (A - \lambda_2 I)x = \begin{bmatrix} 2 & -5 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

El segundo vector característico es cualquier múltiplo diferente de cero de x_2:

$$\textbf{Vector característico para } \lambda_2 \qquad x_2 = \begin{bmatrix} 5 \\ 2 \end{bmatrix}.$$

Quizá el lector observe que las columnas de $A - \lambda_1 I$ proporcionan x_2, y que las columnas de $A - \lambda_2 I$ son múltiplos de x_1. Este hecho es especial (y útil) para las matrices de 2 por 2.

En el caso de 3 por 3, a menudo igualo una componente de x a 1, y resuelvo $(A - \lambda I)x = 0$ para las otras componentes. Por supuesto, si x es un vector característico, entonces también lo son $7x$ y $-x$. Todos los vectores en el espacio nulo de $A - \lambda I$ (que se conoce como *espacio característico*) cumplen $Ax = \lambda x$. En nuestro ejemplo, los espacios característicos son las rectas que pasan por $x_1 = (1, 1)$ y $x_2 = (5, 2)$.

Antes de volver a la aplicación (la ecuación diferencial), se recalcan los pasos para resolver $Ax = \lambda x$:

1. ***Calcular el determinante de $A - \lambda I$.*** Una vez que λ se ha restado de la diagonal principal, este determinante es un polinomio de grado n. Comienza con $(-\lambda)^n$.
2. ***Encontrar las raíces de este polinomio.*** Las n raíces son los valores característicos de A.
3. ***Para cada valor característico, resolver la ecuación $(A - \lambda I)x = 0$.*** Debido a que el determinante es cero, hay otras soluciones distintas de $x = 0$. Se trata de los vectores característicos.

En la ecuación diferencial, lo anterior produce las soluciones especiales $u = e^{\lambda t}x$. Se trata de las *soluciones puramente exponenciales* de $du/dt = Au$. Observe a e^{-t} y e^{2t}:

$$u(t) = e^{\lambda_1 t}x_1 = e^{-t}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \text{y} \qquad u(t) = e^{\lambda_2 t}x_2 = e^{2t}\begin{bmatrix} 5 \\ 2 \end{bmatrix}.$$

Estas dos soluciones especiales proporcionan la solución completa. Pueden multiplicarse por números cualesquiera c_1 y c_2, y pueden sumarse entre sí. Cuando u_1 y u_2 satisfacen la ecuación lineal $du/dt = Au$, también lo hace su suma $u_1 + u_2$:

$$\textbf{Solución completa} \qquad u(t) = c_1 e^{\lambda_1 t}x_1 + c_2 e^{\lambda_2 t}x_2 \tag{12}$$

Este hecho se denomina ***superposición,*** y es válido para las ecuaciones diferenciales (homogéneas y lineales) justo como era válido para ecuaciones matriciales $Ax = 0$. El espacio nulo siempre es un subespacio, y las combinaciones de las soluciones siguen siendo soluciones.

Ahora se tienen dos parámetros libres c_1 y c_2, y es razonable esperar que sea posible de modo que cumplan la condición inicial $u = u(0)$ en $t = 0$:

$$\textbf{Condición inicial} \qquad c_1 x_1 + c_2 x_2 = u(0) \text{ o bien } \begin{bmatrix} 1 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 5 \end{bmatrix}. \tag{13}$$

Las constantes son $c_1 = 3$ y $c_2 = 1$, y ***la solución de la ecuación original es***

$$u(t) = 3e^{-t}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + e^{2t}\begin{bmatrix} 5 \\ 2 \end{bmatrix}. \tag{14}$$

Cuando las dos componentes se escriben por separado, se tiene $v(0) = 8$ y $w(0) = 5$:

$$\textbf{Solución} \qquad v(t) = 3e^{-t} + 5e^{2t}, \qquad w(t) = 3e^{-t} + 2e^{2t}.$$

La clave se encontraba en los valores característicos λ y en los vectores característicos x. Los valores característicos son importantes de suyo, y no sólo son parte de un truco para encontrar u. Quizá el ejemplo más conocido es el de los soldados marchando sobre un puente.* Tradicionalmente, dejan de marchar y simplemente caminan sobre el puente. Si ocurre que caminan a una frecuencia igual a uno de los valores característicos del puente, entonces éste comienza a oscilar. (Del mismo modo en que lo hace un columpio infantil; pronto se percibe la frecuencia natural de un giro, e igualándola se hace más grande la oscilación). Un ingeniero intenta que las frecuencias naturales de su puente o nave espacial estén alejadas de las frecuencias del viento o del movimiento del combustible. Y en el otro extremo, un corredor de bolsa pasa su vida intentando estar en línea con las frecuencias naturales del mercado. Los valores característicos constituyen el rasgo más importante de prácticamente cualquier sistema dinámico.

Resumen y ejemplos

Para resumir, en esta introducción se ha mostrado cómo λ y x aparecen de forma natural y espontánea cuando se resuelve $du/dt = Au$. Esta ecuación posee *soluciones puramente exponenciales* $u = e^{\lambda t}x$; el valor característico proporciona la razón de crecimiento o disminución, y el vector característico x se desarrolla a esta razón. Las otras soluciones son *mezclas* de estas soluciones puras, y la mezcla se ajusta para cumplir las condiciones iniciales.

La ecuación clave era $Ax = \lambda x$. La mayor parte de los vectores no satisfacen esta ecuación. Cambian de dirección cuando se multiplican por A, de modo que Ax no es múltiplo de x. Esto significa que ***sólo ciertos números especiales λ son valores característicos, y sólo ciertos vectores especiales x son vectores característicos.*** Es posible observar el com-

*Ejemplo en el que nunca creí, aunque un puente se cayó de esta forma en 1831.

portamiento de cada vector característico, y luego combinar estos "modos normales" para encontrar la solución. En otras palabras, *es posible diagonalizar la matriz subyacente.*

La diagonalización de la sección 5.2 se aplicará a ecuaciones en diferencias, a los números de Fibonacci, y a los procesos de Markov, así como a las ecuaciones diferenciales. En cada ejemplo se comenzará con el cálculo de los valores característicos y los vectores característicos; no existe ningún atajo para evitar esto. Las matrices simétricas son especialmente fáciles. Las "matrices defectuosas" carecen de un conjunto completo de vectores característicos, por lo que no son diagonalizables. Ciertamente, es necesario analizarlas, aunque no se les permitirá su ingreso en el libro.

Se empieza con ejemplos de matrices especialmente bondadosas.

Ejemplo 1 Todo resulta evidente cuando A es una ***matriz diagonalizable***:

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{tiene} \quad \lambda_1 = 3 \quad \text{con} \quad x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \lambda_2 = 2 \quad \text{con} \quad x_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Sobre cada vector característico, A actúa como un múltiplo de la identidad: $Ax_1 = 3x_1$ y $Ax_2 = 2x_2$. Otros vectores como $x = (1, 5)$ son mezclas de $x_1 + 5x_2$ de los dos vectores característicos, y cuando A multiplica a x_1 y x_2, se obtienen los valores característicos $\lambda_1 = 3$ y $\lambda_2 = 2$:

$$A \text{ multiplicada por} \quad x_1 + 5x_2 \quad \text{es} \quad 3x_1 + 10x_2 = \begin{bmatrix} 3 \\ 10 \end{bmatrix}.$$

Esto es Ax para un vector típico x, no para un vector característico. Sin embargo, la acción de A es determinada por sus vectores característicos y valores característicos:

Ejemplo 2 ¡Los valores característicos de una ***matriz proyección*** son 1 o 0!

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad \text{tiene} \quad \lambda_1 = 1 \quad \text{con} \quad x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \lambda_2 = 0 \quad \text{con} \quad x_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Cuando x se proyecta sobre sí mismo, se tiene $\lambda = 1$, y cuando x se proyecta sobre el vector cero se tiene $\lambda = 0$. El espacio columna de P está lleno de vectores característicos, así como el espacio nulo. Si las dimensiones de estos espacios son r y $n - r$, respectivamente, entonces $\lambda = 1$ se repite r veces y $\lambda = 0$ se repite $n - r$ veces (*siempre n λs*):

$$\textbf{Cuatro valores característicos que permiten repeticiones} \qquad P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{tiene} \quad \lambda = 1, 1, 0, 0.$$

No hay nada excepcional sobre $\lambda = 0$. Así como cualquier otro número, el cero podría o no ser un valor característico. En caso de serlo, entonces sus vectores característicos satisfacen $Ax = 0x$. Así, x está en el espacio nulo de A. Un valor característico cero indica que A es singular (no invertible); su determinante es cero. Todas las matrices invertibles cumplen $\lambda \neq 0$.

Ejemplo 3 Cuando A es ***triangular***, los valores característicos están sobre la diagonal principal.

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 4 & 5 \\ 0 & \frac{3}{4}-\lambda & 6 \\ 0 & 0 & \frac{1}{2}-\lambda \end{vmatrix} = (1-\lambda)\left(\tfrac{3}{4}-\lambda\right)\left(\tfrac{1}{2}-\lambda\right).$$

El determinante es justo el producto de los elementos en la diagonal. Es cero si $\lambda = 1$, $\lambda = \frac{3}{4}$ o $\lambda = \frac{1}{2}$; los valores característicos ya estaban en la diagonal principal.

Este ejemplo, donde los valores característicos pueden encontrarse por inspección, apunta a otro tema fundamental del capítulo: la transformación de A en una matriz diagonal o triangular *sin cambiar sus valores característicos*. Una vez más se recalca que la factorización gaussiana $A = LU$ no es adecuada para este propósito. Los valores característicos de U pueden ser visibles en la diagonal, pero ***no*** son los valores característicos de A.

Para la mayor parte de las matrices, no hay duda de que el problema de los valores característicos es computacionalmente más difícil que $Ax = b$. Con sistemas lineales, un número finito de pasos de eliminación producía la respuesta exacta en un tiempo finito. (O, de manera equivalente, con la regla de Cramer se obtenía una fórmula exacta para la solución). Ninguna fórmula es capaz de proporcionar los valores característicos, o Galois se revolvería en su tumba. Para una matriz de 5 por 5, det $(A - \lambda I)$ implica λ^5. Galois y Abel demostraron que no puede haber ninguna fórmula algebraica para encontrar las raíces de un polinomio de quinto grado.

Todo lo que éstos permiten son algunas verificaciones sencillas de sus valores característicos, *después* que se han calculado, y se mencionan dos buenos: ***la suma y el producto.***

5B La ***suma*** de los n valores característicos es igual a la suma de los n elementos en la diagonal:

$$\textbf{Traza de} \quad A = \lambda_1 + \cdots + \lambda_n = a_{11} + \cdots + a_{nn}. \tag{15}$$

Además, el ***producto*** de los n valores característicos es igual al ***determinante*** de A.

La matriz proyección P tiene elementos en la diagonal $\frac{1}{2}, \frac{1}{2}$ y valores característicos 1, 0 Así, $\frac{1}{2} + \frac{1}{2}$ coincide con $1 + 0$, como debe ser. Así también lo hace el determinante, que es $0 \cdot 1 = 0$. Una matriz singular, con determinante cero, tiene uno o más de sus valores característicos igual a cero.

No debe haber confusión entre los elementos en la diagonal y los valores característicos. Para una matriz triangular siempre son iguales, aunque este hecho es excepcional. Normalmente los pivotes, los elementos en la diagonal, y los valores característicos son completamente distintos. Y para una matriz de 2 por 2, la traza y el determinante lo dicen todo:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{tiene traza } a + d, \text{ y determinante } ad - bc$$

$$\det(A - \lambda I) = \det \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = \lambda^2 - (\text{traza})\lambda + \text{determinante}$$

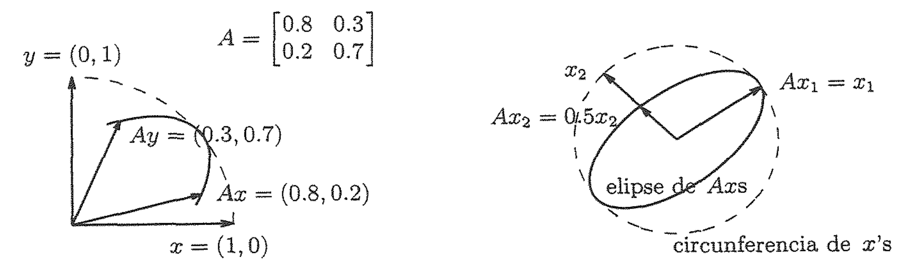
$$\text{Los valores característicos son } \lambda = \frac{\text{traza} \pm [(\text{traza})^2 - 4\det]^{1/2}}{2}.$$

La suma de estos dos λs es igual a la traza; en el ejercicio 9 se proporciona $\sum \lambda_i = \text{traza}$ para todas las matrices.

Eigshow

Hay un programa de demostración MATLAB (simplemente hay que teclear eigshow), que despliega el problema del valor característico para una matriz de 2 por 2. Empieza con el vector unitario $x = (1, 0)$. *El mouse hace que este vector se desplace alrededor de la circun-*

ferencia unitaria. Al mismo tiempo, la pantalla muestra a Ax, a color y también en movimiento. Posiblemente Ax esté enfrente de x. Posiblemente Ax esté detrás de x. *Algunas veces Ax es paralelo a x.* En ese instante paralelo, $Ax = \lambda x$ (dos veces en la segunda figura).



El valor característico λ es la longitud de Ax, cuando el vector característico unitario x es paralelo. Las opciones integradas para A, ilustran tres posibilidades: 0, 1, o 2 para vectores característicos reales.

1. No hay *vectores característicos reales. Ax permanece atrás o adelante de x.* Esto significa que los valores característicos y los vectores característicos, son complejos, como lo son para la rotación Q.
2. Hay *sólo una* recta de vectores característicos (lo cual es inusual). Las direcciones cambiantes Ax y x se encuentran pero no se cortan. Este hecho ocurre para la siguiente matriz de 2 por 2.
3. Hay vectores característicos en *dos* direcciones independientes. ¡Este hecho es típico! Ax corta a x en el primer vector característico x_1, y corta en el segundo vector característico x_2.

Suponga que A es singular (de rango 1). Su espacio columna es una recta. El vector Ax debe permanecer en esa recta mientras x gira alrededor. Un vector característico x está a lo largo de la recta. Otro vector característico aparece cuando $Ax_2 = 0$. Cero es un valor característico de una matriz singular.

Para estas seis matrices, es posible seguir mentalmente a x y a Ax. ¿Cuántos vectores característicos hay, y dónde? ¿Cuándo ocurre que Ax se desplaza en el sentido del movimiento de las manecillas del reloj, en vez de hacerlo en sentido contrario a las manecillas del reloj con x?

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Conjunto de problemas 5.1

1. Encuentre los valores característicos y los vectores característicos de la matriz $A = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$. Compruebe que la traza es igual a la suma de los valores característicos, y que el determinante es igual a su producto.

2. Con la misma matriz A, resuelva la ecuación diferencial $du/dt = Au$, $u(0) = \begin{bmatrix} 0 \\ 6 \end{bmatrix}$. ¿Cuáles son las dos soluciones exponenciales puras?

3. Si se pasa a $A - 7I$, ¿cuáles son los valores característicos y los vectores característicos, y cómo están relacionados con los de A?

$$B = A - 7I = \begin{bmatrix} -6 & -1 \\ 2 & -3 \end{bmatrix}.$$

4. Resuelva $du/dt = Pu$ cuando P es una proyección:

$$\frac{du}{dt} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} u \quad \text{con} \quad u(0) = \begin{bmatrix} 5 \\ 3 \end{bmatrix}.$$

Parte de $u(0)$ crece exponencialmente mientras la parte del espacio nulo, permanece fija.

5. Encuentre los valores característicos y los vectores característicos de

$$A = \begin{bmatrix} 3 & 4 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{bmatrix}.$$

Compruebe que $\lambda_1 + \lambda_2 + \lambda_3$ es igual a la traza y que $\lambda_1\lambda_2\lambda_3$ es igual al determinante.

6. Proporcione un ejemplo para demostrar que los valores característicos pueden cambiarse cuando un múltiplo de un renglón se resta de otro. ¿Por qué los pasos de eliminación *no* modifican un valor característico igual a cero?

7. Suponga que λ es un valor característico de A, y que x es un vector característico: $Ax = \lambda x$:

   a) Demuestre que este mismo x es un vector característico de $B = A - 7I$, y encuentre el valor característico. Esto debe confirmar el ejercicio 3.

   b) Suponga que $\lambda \neq 0$, demuestre que x también es un vector característico de A^{-1}, y encuentre el valor característico.

8. Demuestre que el determinante es igual al producto de los valores característicos, suponiendo que el polinomio característico se factoriza como

$$\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda), \tag{16}$$

y haciendo una elección inteligente de λ.

9. En dos pasos, demuestre que la traza es igual a la suma de los valores característicos. Primero, encuentre el coeficiente de $(-\lambda)^{n-1}$ en el miembro derecho de la ecuación (16). Luego, encuentre todos los términos de

$$\det(A - \lambda I) = \det \begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{bmatrix}$$

que implican a $(-\lambda)^{n-1}$. ¡Todos provienen de la diagonal principal! Encuentre ese coeficiente de $(-\lambda)^{n-1}$, y compare.

10. *a*) Construya matrices de 2 por 2 tales que los valores característicos de $A\,B$ no sean los productos de los valores característicos de A y B, y los valores característicos de $A + B$ no sean las sumas de los valores característicos individuales.

    b) Compruebe, no obstante, que la suma de los valores característicos de $A + B$ es igual a la suma de todos los valores característicos individuales de A y B, y de manera semejante para los productos. ¿Por qué es cierto lo anterior?

11. ***Los valores característicos de A son iguales a los valores característicos de A^T.*** Esto se debe a que $\det(A - \lambda I)$ es igual a $\det(A^T - \lambda I)$. Lo anterior es cierto porque ______. Demuestre con un ejemplo que los vectores característicos de A y A^T *no* son los mismos.

12. Encuentre los valores característicos y los vectores característicos de

$$A = \begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}.$$

13. Si los valores característicos de B son 1, 2, 3, los valores característicos de C son 4, 5, 6, y los valores característicos de D son 7, 8, 9, ¿cuáles son los valores característicos de la matriz de 6 por 6 $A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$?

14. Encuentre el rango y todos los cuatro valores característicos de la matriz de 1s y la matriz de tablero de ajedrez:

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

¿Cuáles vectores característicos corresponden a valores característicos diferentes de cero?

15. ¿Cuáles son el rango y los valores característicos cuando A y C en el ejercicio previo son de n por n? Recuerde que el valor característico $\lambda = 0$ se repite $n - r$ veces.

16. Si A es la matriz de 4 por 4 de 1s, encuentre los valores característicos y el determinante de $A - I$.

17. Escoja el tercer renglón de la "matriz siguiente"

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \cdot & \cdot & \cdot \end{bmatrix}$$

de modo que su polinomio característico $|A - \lambda I|$ sea $-\lambda^3 + 4\lambda^2 + 5\lambda + 6$.

18. Suponga que los valores característicos de A son 0, 3, 5, con vectores característicos independientes u, v, w.

a) Proporcione una base para el espacio nulo y una base para el espacio columna.
b) Encuentre una solución particular de $Ax = v + w$. Encuentre todas las soluciones.
c) Demuestre que $Ax = u$ no tiene solución. (En caso de tenerla, entonces ______ estaría en el espacio columna.)

19. Las potencias A^k de esta matriz A tienden al límite cuando $k \to \infty$:

$$A = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 0.70 & 0.45 \\ 0.30 & 0.55 \end{bmatrix}, \quad \text{y} \quad A^\infty = \begin{bmatrix} 0.6 & 0.6 \\ 0.4 & 0.4 \end{bmatrix}.$$

La matriz A^2 está a mitad del camino entre A y A^∞. Explique por qué $A^2 = \frac{1}{2}(A + A^\infty)$, a partir de los valores característicos y vectores característicos de estas tres matrices.

20. Encuentre los valores característicos y los vectores característicos de las dos siguientes matrices:

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} \quad \text{y} \quad A + I = \begin{bmatrix} 2 & 4 \\ 2 & 4 \end{bmatrix}.$$

$A + I$ tiene los ____ vectores característicos que A. Sus valores característicos están ______ por 1.

21. Calcule los valores característicos y los vectores característicos de A y A^{-1}:

$$A = \begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix} \quad \text{y} \quad A^{-1} = \begin{bmatrix} -3/4 & 1/2 \\ 1/2 & 0 \end{bmatrix}.$$

A^{-1} tiene los ____ vectores característicos que A. Cuando los valores característicos de A son λ_1 y λ_2, los valores característicos de su inversa son _____.

22. Calcule los valores característicos y los vectores característicos de A y A^2:

$$A = \begin{bmatrix} -1 & 3 \\ 2 & 0 \end{bmatrix} \quad \text{y} \quad A^2 = \begin{bmatrix} 7 & -3 \\ -2 & 6 \end{bmatrix}.$$

A^2 tiene los mismos____ que A. Cuando los valores característicos de A son λ_1 y λ_2, los valores característicos de A^2 son _____.

23. *a*) Si se sabe que x es un vector característico, la forma de encontrar λ es _____.
b) Si se sabe que λ es un valor característico, la forma de encontrar x es _____.

24. ¿Qué se hace a $Ax = \lambda x$, para demostrar los incisos *a*), *b*), y *c*)?
a) λ^2 es un valor característico de A^2, como en el problema 22.
b) λ^{-1} es un valor característico de A^{-1}, como en el problema 21.
c) $\lambda + 1$ es un valor característico de $A + I$, como en el problema 20.

25. A partir del vector unitario $u = \left(\frac{1}{6}, \frac{1}{6}, \frac{3}{6}, \frac{5}{6}\right)$, construya la matriz proyección de rango 1 $P = uu^{\mathrm{T}}$.
a) Demuestre que $Pu = u$. Así, u es un vector característico con $\lambda = 1$.
b) Si v es perpendicular a u, demuestre que Pv = vector cero. Así, $\lambda = 0$.
c) Encuentre tres vectores característicos independientes de P, todos con valor característico $\lambda = 0$.

26. Resuelva $\det(Q - \lambda I) = 0$, aplicando la fórmula cuadrática, para llegar a $\lambda = \cos\theta \pm i \operatorname{sen}\theta$:

$$Q = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \quad \text{rota el plano } x\text{-}y \text{ por el ángulo } \theta.$$

Resuelva $(Q - \lambda I)x = 0$ para encontrar los vectores característicos de Q. Use $i^2 = -1$.

27. Toda matriz permutación deja sin cambio a $x = (1, 1, \ldots, 1)$. Así, $\lambda = 1$. Encuentre dos λs más para las siguientes permutaciones:

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{y} \quad P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

28. Si A tiene $\lambda_1 = 4$ y $\lambda_2 = 5$, entonces $\det(A - \lambda I) = (\lambda - 4)(\lambda - 5) = \lambda^2 - 9\lambda + 20$. Encuentre tres matrices que cumplan lo siguiente: traza $a + d = 9$, determinante = 20, y $\lambda = 4, 5$.

29. Se sabe que los valores característicos de una matriz B de 3 por 3 son 0, 1, 2. Esta información es suficiente para encontrar tres de los cuatro incisos siguientes:
a) el rango de A,
b) el determinante de $B^{\mathrm{T}}B$,

c) los valores característicos de B^TB, y

d) los valores característicos de $(B + I)^{-1}$.

30. Escoja el segundo renglón de $A = \begin{bmatrix} 0 & 1 \\ * & * \end{bmatrix}$ de modo que los valores característicos de A sean 4 y 7.

31. Escoja a, b, c, de modo que $\det(A - \lambda I) = 9\lambda - \lambda^3$. Así, los valores característicos son $-3, 0, 3$:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a & b & c \end{bmatrix}.$$

32. Construya una matriz M de Markov de 3 por 3: la suma de los elementos positivos de cada columna es 1. Si $e = (1, 1, 1)$, compruebe que $M^Te = e$. Por el problema 11, $\lambda = 1$ también es un valor característico de M. Desafío: Una matriz singular M de Markov de 3 por 3 con traza $\frac{1}{2}$ tiene valores característicos $\lambda =$ ______.

33. Encuentre tres matrices de 2 por 2 que tengan $\lambda_1 = \lambda_2 = 0$. La traza es cero y el determinante es cero. La matriz A podría no ser 0, pero compruebe que $A^2 = 0$.

34. La siguiente matriz es singular con rango 1. Encuentre tres λs y tres vectores característicos:

$$A = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 2 \\ 4 & 2 & 4 \\ 2 & 1 & 2 \end{bmatrix}.$$

35. Suponga que A y B tienen los mismos valores característicos $\lambda_1, \ldots, \lambda_n$ con los mismos vectores característicos independientes $x_1, \ldots, x_n$. Así, $A = B$. *Razón*: Cualquier vector x es una combinación $c_1x_1 + \cdots + c_nx_n$. ¿Cuál es Ax? ¿Cuál es Bx?

36. (Repaso) Encuentre los valores característicos de A, B, y C:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 3 & 0 & 0 \end{bmatrix}, \quad \text{y} \quad C = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}.$$

37. Cuando $a + b = c + d$, demuestre que $(1, 1)$ es un vector característico, y encuentre ambos valores característicos:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

38. Cuando P intercambia los renglones 1 y 2 y las columnas 1 y 2, los valores característicos no cambian. Encuentre vectores característicos de A y PAP para $\lambda = 11$:

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 6 & 3 \\ 4 & 8 & 4 \end{bmatrix} \quad \text{y} \quad PAP = \begin{bmatrix} 6 & 3 & 3 \\ 2 & 1 & 1 \\ 8 & 4 & 4 \end{bmatrix}.$$

39. Problema de desafío: *¿Existe alguna matriz real de* 2 *por* 2 (distinta de I) *con* $A^3 = I$? Sus valores característicos deben cumplir $\lambda^3 = I$. Pueden ser $e^{2\pi i/3}$ y $e^{-2\pi i/3}$. ¿Qué traza y qué determinante se obtienen con lo anterior? Construya A.

40. Hay seis matrices permutación P de 3 por 3. ¿Cuáles números pueden ser los *determinantes* de P? ¿Qué números pueden ser los *pivotes*? ¿Qué números pueden ser la *traza* de P? ¿Cuáles *cuatro números* pueden ser valores característicos de P?

5.2 DIAGONALIZACIÓN DE UNA MATRIZ

Se empieza directamente con el cálculo esencial. Es perfectamente sencillo, y se utilizará en todas las secciones de este capítulo. ***Los vectores característicos diagonalizan una matriz:***

5C Suponga que la matriz A de n por n tiene n vectores característicos linealmente independientes. Si estos vectores característicos son las columnas de una matriz S, entonces $S^{-1}AS$ es una matriz diagonal Λ. Los valores característicos de A están sobre la diagonal de Λ:

Diagonalización
$$S^{-1}AS = \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}. \qquad (1)$$

La matriz S se denomina "matriz vector característico", y Λ se denomina "matriz valor característico", para la cual se utiliza una letra lambda mayúscula, ya que las lambdas minúsculas se usan para designar los valores característicos que están en su diagonal.

Demostración Los vectores característicos x_i se escriben en las columnas de S, y AS se calcula por columnas:

$$AS = A \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \lambda_1x_1 & \lambda_2x_2 & \cdots & \lambda_nx_n \\ | & | & & | \end{bmatrix}.$$

Luego, el truco consiste en separar esta última matriz, en un producto $S\Lambda$ bastante diferente:

$$\begin{bmatrix} \lambda_1x_1 & \lambda_2x_2 & \cdots & \lambda_nx_n \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

Es vital mantener estas matrices en el orden correcto. Si Λ está antes de S (y no después), entonces λ_1 multiplicará los elementos de la primera línea. Deseamos que λ_1 aparezca en la primera columna. Si es así, entonces $S\Lambda$ es correcto. Por tanto,

$$AS = S\Lambda, \quad \text{o bien,} \quad S^{-1}AS = \Lambda, \quad \text{o bien,} \quad A = S\Lambda S^{-1}. \qquad (2)$$

S es invertible, ya que se supuso que sus columnas (los vectores característicos) son independientes. Antes de proporcionar ejemplos o aplicaciones es necesario presentar cuatro observaciones. ■

Observación 1 Si la matriz A no tiene valores característicos repetidos; es decir, que los números $\lambda_1, \ldots, \lambda_n$ son distintos, entonces los n vectores característicos son automáticamente independientes (consulte el punto 5D a continuación). En consecuencia, ***cualquier matriz con valores característicos distintos puede diagonalizarse.***

Observación 2 La matriz de diagonalización S *no es única*. Un vector característico x puede multiplicarse por una constante y seguir siendo un vector característico. Las columnas de S pueden multiplicarse por constantes cualesquiera diferentes de cero, y producir

una nueva matriz de diagonalización S. Valores característicos repetidos dejan aún más libertad en S. Para el ejemplo trivial $A = I$, cualquier S logra lo anterior: $S^{-1}IS$ siempre es diagonal (Λ es justo I). Todos los vectores son vectores característicos de la identidad.

Observación 3 *Otras matrices S no producen una* Λ *diagonal.* Suponga que la primera columna de S es y. Entonces la primera columna de $S\Lambda$ es $\lambda_1 y$. Si esto ha de coincidir con la primera columna de AS, que por multiplicación de matrices es Ay, entonces y debe ser un vector característico: $Ay = \lambda_1 y$. El *orden* de los vectores característicos en S y de los valores característicos en Λ es automáticamente el mismo.

Observación 4 No todas las matrices poseen n vectores característicos linealmente independientes, de modo que ***no todas las matrices son diagonalizables.*** El ejemplo estándar de una "matriz defectuosa" es

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Sus valores característicos son $\lambda_1 = \lambda_2 = 0$, ya que la matriz es triangular con ceros en la diagonal:

$$\det(A - \lambda I) = \det \begin{bmatrix} -\lambda & 1 \\ 0 & -\lambda \end{bmatrix} = \lambda^2.$$

Todos los vectores característicos de esta A son múltiplos del vector $(1, 0)$:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \text{o bien,} \quad x = \begin{bmatrix} c \\ 0 \end{bmatrix}.$$

$\lambda = 0$ es un valor característico doble: su *multiplicidad algebraica* es 2. Sin embargo, la *multiplicidad geométrica* es 1, porque sólo hay un vector característico independiente. No es posible construir S.

A continuación se proporciona una demostración más directa de que A no es diagonalizable. Debido a que $\lambda_1 = \lambda_2 = 0$, Λ tendría que ser la matriz cero. Pero si $\Lambda = S^{-1}AS$, $1 = 0$, entonces se premultiplica por la izquierda por S y se posmultiplica por la derecha por S^{-1}, con la finalidad de concluir falsamente que $A = 0$. No hay ninguna S invertible.

Este fracaso de diagonalización ***no*** se debe a que $\lambda = 0$. Proviene del hecho de que $\lambda_1 = \lambda_2$:

$$\textbf{Valores característicos repetidos} \quad A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}.$$

Sus valores característicos son 3, 3 y 1, 1. ¡No son singulares! El problema es la falta de vectores característicos, que se necesitan para S. Es necesario recalcar estas necesidades:

> ***La diagonalización de A depende de la existencia de suficientes vectores característicos.***
> ***La invertibilidad de A depende de los valores característicos diferentes de cero.***

No hay relación entre la diagonalización (n vectores característicos independientes) y la invertibilidad (valores característicos diferentes de cero). La única indicación proporcionada por los valores característicos es: *la diagonalización puede fracasar sólo si hay valores característicos repetidos.* Incluso así, no siempre fracasa. $A = I$ tiene valores característicos repetidos $1, 1, \ldots, 1$, ¡aunque ya es diagonal! En este caso no hay carencia de vectores característicos.

La prueba consiste en comprobar, para un valor característico que se repite p veces, si hay p vectores característicos independientes; en otras palabras, si el rango de $A - \lambda I$ es $n - p$. Para completar este círculo de ideas, es necesario demostrar que valores característicos *distintos* no constituyen ningún problema.

> **5D** Si los vectores característicos $x_1, \ldots, x_k$ corresponde a *valores característicos distintos* $\lambda_1, \ldots, \lambda_k$, entonces estos vectores característicos son linealmente independientes.

Primero se supone que $k = 2$, y que con alguna combinación de x_1 y x_2 se obtiene cero: $c_1x_1 + c_2x_2 = 0$. Al multiplicar por A, se encuentra $c_1\lambda_1x_1 + c_2\lambda_2x_2 = 0$. Al restar λ_2 multiplicado por la ecuación previa, el vector x_2 desaparece:

$$c_1(\lambda_1 - \lambda_2)x_1 = 0.$$

Debido a que $\lambda_1 \neq \lambda_2$ y a que $x_1 \neq 0$, se está forzando a que $c_1 = 0$. De manera semejante, $c_2 = 0$, y los dos vectores son independientes; sólo la combinación trivial proporciona cero.

Este mismo razonamiento se extiende a cualquier cantidad de vectores característicos: si con alguna combinación se obtiene cero, debe multiplicarse por A, restar λ_k multiplicado por la combinación original, y entonces desaparece x_k, dejando una combinación de $x_1, \ldots, x_{k-1}$, con lo cual se obtiene cero. Al repetir los mismos pasos (en realidad, esto es *inducción matemática*) se termina con un múltiplo de x_1 que produce cero. Así se obliga a que $c_1 = 0$, y finalmente a que todo $c_i = 0$. En consecuencia, los vectores característicos que provienen de valores característicos distintos, son automáticamente independientes.

Una matriz con n valores característicos distintos puede diagonalizarse. Este es el caso típico.

Ejemplos de diagonalización

El punto más importante de esta sección es $S^{-1}AS = \Lambda$. La matriz vector característico S transforma A en su matriz valor característico Λ (diagonal). Esto se verá a continuación para proyecciones y rotaciones.

Ejemplo 1 La proyección $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ tiene matriz valor característico $\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Los vectores característicos van en las columnas de S:

$$S = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \text{y} \quad AS = S\Lambda = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

La última ecuación puede comprobarse a primera vista. En consecuencia, $S^{-1}AS = \Lambda$.

Ejemplo 2 Por sí mismos, los valores característicos no resultan tan evidentes para una ***rotación***:

$$\textbf{rotación de 90º} \quad K = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{tiene } \det(K - \lambda I) = \lambda^2 + 1.$$

¿Cómo es posible que un vector rote, y aun así, preserve sin cambio su dirección? Aparentemente, lo anterior no es posible salvo para el vector cero, lo cual es inútil. Sin embargo, debe haber valores característicos, y debe ser posible resolver $du/dt = Ku$. El polinomio característico $\lambda^2 + 1$ debe seguir teniendo dos raíces, aunque *no son reales*.

Ahora, el lector puede ver el camino de salida. Los valores característicos de K son *números imaginarios*, $\lambda_1 = i$ y $\lambda_2 = -i$. Tampoco los vectores característicos son reales.

De alguna manera, al rotar 90°, se multiplican por i o por $-i$:

$$(K - \lambda_1 I)x_1 = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{y} \quad x_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

$$(K - \lambda_2 I)x_2 = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{y} \quad x_2 = \begin{bmatrix} 1 \\ i \end{bmatrix}.$$

Los valores característicos son distintos, aun cuando son imaginarios, y los vectores característicos son independientes. Éstos van en las columnas de S:

$$S = \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} \quad \text{y} \quad S^{-1}KS = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

Se está frente a un hecho ineludible: ***los números complejos son necesarios incluso para matrices reales***. Si hay muy pocos valores característicos reales, siempre hay n valores característicos complejos. (Los complejos incluyen a los reales, cuando la parte imaginaria es cero). Si hay muy pocos vectores característicos en el mundo real $\mathbf{R}^3$, o en $\mathbf{R}^n$, se busca en $\mathbf{C}^3$ o en $\mathbf{C}^n$. El espacio $\mathbf{C}^n$ contiene a todos los vectores columna con componentes complejas, y posee nuevas definiciones de longitud, producto interno y ortogonalidad. Sin embargo, no es más difícil que $\mathbf{R}^n$, y en la sección 5.5 se realiza una sencilla conversión al caso complejo.

Potencias y productos: A^k y AB

Hay una situación adicional en que los cálculos son fáciles. ***Los valores característicos de A^2 son exactamente $\lambda_1^2, \ldots, \lambda_n^2$, y cada vector característico de A también es un vector característico de A^2.*** Se empieza con $Ax = \lambda x$, y se multiplica de nuevo por A:

$$A^2x = A\lambda x = \lambda Ax = \lambda^2 x. \tag{3}$$

Por tanto, λ^2 es un valor característico de A^2, con el mismo vector característico x. Si la primera multiplicación por A deja sin cambio la dirección de x, entonces también lo hace la segunda.

El mismo resultado se obtiene de la diagonalización, al elevar al cuadrado a $S^{-1}AS = \Lambda$:

Valores característicos de A^2 $(S^{-1}AS)(S^{-1}AS) = \Lambda^2$ o bien, $S^{-1}A^2S = \Lambda^2$.

La matriz A^2 es diagonalizada por la misma S, de modo que los vectores característicos permanecen sin cambio. Los valores característicos están al cuadrado. Este hecho se cumple para cualquier potencia de A:

5E Los valores característicos de A^k son $\lambda_1^k, \ldots, \lambda_n^k$, y cada vector característico de A sigue siendo un vector característico de A^k. Una vez que S diagonaliza a A, también diagonaliza a A^k:

$$\Lambda^k = (S^{-1}AS)(S^{-1}AS)\cdots(S^{-1}AS) = S^{-1}A^kS. \tag{4}$$

Cada S^{-1} cancela a una S, excepto por la primera S^{-1} y la última S.

Si A es invertible, esta regla también es válida para su inversa (la potencia $k = -1$). ***Los valores característicos de A^{-1} son $1/\lambda_i$.*** Este hecho puede verse incluso, sin diagonalizar:

$$\text{Si} \quad Ax = \lambda x \quad \text{entonces} \quad x = \lambda A^{-1}x \quad \text{y} \quad \frac{1}{\lambda}x = A^{-1}x.$$

Ejemplo 3 Si K es una rotación de 90°, entonces K^2 es una rotación de 180° (lo cual significa $-I$) y K^{-1} es una rotación de $-90°$:

$$K = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad K^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \text{y} \quad K^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Los valores característicos de K son i y $-i$; sus cuadrados son -1 y -1; sus recíprocos son $1/i = -i$ y $1/(-i) = i$. Así, K^4 es una rotación completa de 360°:

$$K^4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y también} \quad \Lambda^4 = \begin{bmatrix} i^4 & 0 \\ 0 & (-i)^4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Para un ***producto de dos matrices***, puede preguntarse sobre los valores característicos de AB, aunque no se obtiene una buena respuesta. Es bastante tentador tratar de aplicar el mismo razonamiento, esperando demostrar que *en general no es cierto*. Si λ es un valor característico de A y μ es un valor característico de B, entonces he aquí la falsa demostración de que AB tiene el valor característico, $\mu\lambda$:

Demostración falsa $ABx = A\mu x = \mu Ax = \mu\lambda x.$

El error consiste en suponer que A y B comparten *el mismo* vector característico x. En general, no es así. Es posible que haya dos matrices con valores característicos cero, mientras AB tiene valor característico $\lambda = 1$:

$$AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Los vectores característicos de estas A y B son totalmente distintos, lo cual es típico. Por la misma razón, en general los valores característicos de $A + B$ no tienen nada que ver con $\lambda + \mu$.

Esta falsa demostración sugiere *ser verdadera*. Si el vector característico es el mismo para A y B, entonces los valores característicos se multiplican y el valor característico de AB es $\mu\lambda$. Sin embargo, se tiene algo más importante. Hay una forma fácil de identificar cuándo A y B comparten un conjunto completo de vectores característicos, lo cual plantea una cuestión clave en mecánica cuántica:

5F Las matrices diagonalizables comparten la misma matriz vector característico S si y sólo si $AB = BA$.

Demostración Si la misma S diagonaliza tanto a $A = S\Lambda_1S^{-1}$ y $B = S\Lambda_2S^{-1}$, puede multiplicarse en cualquier orden:

$$AB = S\Lambda_1S^{-1}S\Lambda_2S^{-1} = S\Lambda_1\Lambda_2S^{-1} \quad \text{y} \quad BA = S\Lambda_2S^{-1}S\Lambda_1S^{-1} = S\Lambda_2\Lambda_1S^{-1}.$$

Debido a que $\Lambda_1\Lambda_2 = \Lambda_2\Lambda_1$ (las matrices diagonales siempre son conmutativas) se tiene que $AB = BA$.

En la dirección opuesta, suponga que $AB = BA$. Empezando con $Ax = \lambda x$, se tiene

$$ABx = BAx = B\lambda x = \lambda Bx.$$

Así, tanto x como Bx son vectores característicos de A, que comparten el mismo λ (o bien, $Bx = 0$). Si por conveniencia se supone que los valores característicos de A son distintos; es decir, que todos los espacios característicos son unidimensionales, entonces *Bx debe ser un múltiplo de x*. En otras palabras, x es un vector característico de B, así como de A. La demostración con valores característicos repetidos es un poco más larga. ■

El ***principio de incertidumbre de Heisenberg*** proviene de las matrices no conmutativas, como la posición P y la cantidad de movimiento Q. La posición es simétrica, la cantidad de movimiento es simétrica sesgada, y juntas cumplen $QP - PQ = I$. El principio de incertidumbre se concluye directamente de la desigualdad de Schwarz $(Qx)^{\mathrm{T}}(Px) \leq \|Qx\|\|Px\|$ de la sección 3.2:

$$\|x\|^2 = x^{\mathrm{T}}x = x^{\mathrm{T}}(QP - PQ)x \leq 2\|Qx\|\|Px\|.$$

El producto de $\|Qx\|/\|x\|$ y $\|Px\|/\|x\|$ —los errores de posición y cantidad de movimiento, cuando la función de onda es x— es por lo menos $\frac{1}{2}$. Es imposible que ambos errores sean pequeños, porque al intentar medir la posición de una partícula se modifica su cantidad de movimiento.

Al final se volverá a $A = S\Lambda S^{-1}$. Esta factorización es particularmente idónea para tomar potencias de A, y la cuestión se ilustra con el caso más simple: A^2. La factorización LU es inútil cuando se eleva al cuadrado, pero $S\Lambda S^{-1}$ es perfecta. El cuadrado es $S\Lambda^2 S^{-1}$, y los vectores característicos permanecen sin cambio. Trabajar con estos vectores característicos permite resolver ecuaciones en diferencias y ecuaciones diferenciales.

Conjunto de problemas 5.2

1. Factorice las siguientes matrices en $S\Lambda S^{-1}$:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}.$$

2. Encuentre la matriz A cuyos valores característicos son 1 y 4, y cuyos vectores característicos son $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, respectivamente. (*Sugerencia*: $A = S\Lambda S^{-1}$.)

3. Encuentre *todos* los valores característicos y los vectores característicos de

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

y escriba dos matrices de diagonalización S distintas.

4. Si los elementos diagonales de una matriz triangular superior de 3 por 3 son 1, 2, 7, ¿cómo puede saberse que la matriz es diagonalizable? ¿Cuál es Λ?

5. De las siguientes matrices, ¿cuál(es) no puede(n) diagonalizarse?

$$A_1 = \begin{bmatrix} 2 & -2 \\ 2 & -2 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 2 & 0 \\ 2 & -2 \end{bmatrix} \qquad A_3 = \begin{bmatrix} 2 & 0 \\ 2 & 2 \end{bmatrix}.$$

6. *a*) Si $A^2 = I$, ¿cuáles son los posibles valores característicos de A?
b) Si esta matriz A es de 2 por 2, y no es I o $-I$, encuentre su traza y su determinante.
c) Si el primer renglón es $(3, -1)$, ¿cuál es el segundo renglón?

7. Si $A = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$, encuentre A^{100}, diagonalizando A.

8. Suponga que $A = uv^{\mathrm{T}}$ es una columna multiplicada por un renglón (una matriz con rango 1).

a) Multiplique A por u, para demostrar que u es un vector característico. ¿Cuál es λ?
b) ¿Cuáles son los otros valores característicos de A (y por qué)?
c) Calcule traza (A), a partir de la suma de la diagonal y la suma de los λs.

9. Demuestre por cálculo directo que AB y BA tienen la misma traza cuando

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} q & r \\ s & t \end{bmatrix}.$$

Concluya que $AB - BA = I$ es imposible (excepto en dimensiones infinitas).

10. Suponga que los valores característicos de A son 1, 2, 4. ¿Cuál es la traza de A^2? ¿Cuál es el determinante de $(A^{-1})^{\mathrm{T}}$?

11. Si los valores característicos de A son 1, 1, 2, ¿de cuáles de las siguientes afirmaciones se tiene la certeza de que son verdaderas?
Proporcione un razonamiento de por qué sí son verdaderas o un ejemplo si no son verdaderas.

a) A es invertible.
b) A es diagonalizable.
c) A no es diagonalizable.

12. Suponga que sólo los vectores característicos de A son múltiplos de $x = (1, 0, 0)$. Falso o verdadero:
a) A no es invertible.
b) A tiene un valor característico repetido.
c) A no es diagonalizable.

13. Diagonalice la matriz $A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$ y encuentre una de sus raíces cuadradas: una matriz tal que $R^2 = A$. ¿Cuántas raíces cuadradas hay?

14. Suponga que la matriz vector característico S tiene $S^{\mathrm{T}} = S^{-1}$. Demuestre que $A = S\Lambda S^{-1}$ es simétrica y tiene vectores característicos ortogonales.

Los problemas 15 a 24 son sobre matrices valor característico, y vector característico.

15. Factorice las siguientes matrices en $A = S\Lambda S^{-1}$:

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}.$$

16. Si $A = S\Lambda S^{-1}$ entonces $A^3 = (\quad)(\quad)(\quad)$ y $A^{-1} = (\quad)(\quad)(\quad)$.

17. Si A tiene $\lambda_1 = 2$ con vector característico $x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\lambda_2 = 5$ con $x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, use $S\Lambda S^{-1}$ para encontrar A. Ninguna otra matriz tiene los mismos λs y xs.

18. Suponga que $A = S\Lambda S^{-1}$. ¿Cuál es la matriz valor característico para $A + 2I$? ¿Cuál es la matriz vector característico? Compruebe que $A + 2I = (\quad)(\quad)(\quad)^{-1}$.

19. ¿Falso o verdadero? Si las n columnas de S (vectores característicos de A) son independientes, entonces
a) A es invertible.
b) A es diagonalizable.
c) S es invertible.
d) S es diagonalizable.

20. Si los vectores característicos de A son las columnas de I, entonces A es una matriz ______. Si la matriz vector característico S es triangular, entonces S^{-1} es triangular y A es triangular.

21. Describa todas las matrices S que diagonalizan la siguiente matriz A:

$$A = \begin{bmatrix} 4 & 0 \\ 1 & 2 \end{bmatrix}.$$

Luego describa todas las matrices que diagonalizan a A^{-1}.

22. Escriba la matriz más general que tenga vectores característicos $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

23. Encuentre los valores característicos de A y B y $A + B$:

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad A + B = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Los valores característicos de $A + B$ (son iguales a)(no son iguales a) los valores característicos de A más los valores característicos de B.

24. Encuentre los valores característicos de A, B, AB y BA:

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad AB = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad \text{y} \quad BA = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

Los valores característicos de AB (son iguales a)(no son iguales a) los valores característicos de A multiplicados por los valores característicos de B. Los valores característicos de $A\,B$ (son)(no son) iguales a los valores característicos de BA.

Los problemas 25 a 28 son sobre la diagonalizabilidad de A.

25. ¿Falso o verdadero? Si los valores característicos de A son 2, 2, 5, entonces la matriz ciertamente es

a) Invertible.
b) Diagonalizable.
c) No diagonalizable.

26. Si los valores característicos de A son 1 y 0, escriba todo lo que sepa sobre las matrices A y A^2.

27. Complete las siguientes matrices de modo que $\det A = 25$. Así, traza $= 10$, y ¡$\lambda = 5$ es repetido! Encuentre un vector característico con $Ax = 5x$. Estas matrices no son diagonalizables porque no hay una segunda línea de vectores característicos.

$$A = \begin{bmatrix} 8 & \\ & 2 \end{bmatrix}, \quad A = \begin{bmatrix} 9 & 4 \\ & 1 \end{bmatrix}, \quad \text{y} \quad A = \begin{bmatrix} 10 & 5 \\ -5 & \end{bmatrix}.$$

28. La matriz $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ no es diagonalizable porque el rango de $A - 3I$ es _____.
Cambie un elemento para hacer diagonalizable a A. ¿Qué elementos podría cambiar?

Los problemas 29 a 33 son sobre potencias de matrices.

29. $A^k = S\Lambda^k S^{-1}$ tiende a la matriz cero cuando $k \to \infty$ si y sólo si el valor absoluto de cada λ es menor que _____. ¿Qué es cierto: $A^k \to 0$ o $B^k \to 0$?

$$A = \begin{bmatrix} 0.6 & 0.4 \\ 0.4 & 0.6 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0.6 & 0.9 \\ 0.1 & 0.6 \end{bmatrix}.$$

30. (Recomendado) Encuentre Λ y S para diagonalizar A en el problema 29. ¿Cuál es el límite de Λ^k cuando $k \to \infty$? ¿Cuál es el límite de $S\Lambda^k S^{-1}$? En las columnas de esta matriz limitante, es posible observar el _____.

31. Encuentre Λ y S para diagonalizar B en el problema 29. ¿Cuál es $B^{10}u_0$ para estos u_0?

$$u_0 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad u_0 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \quad \text{y} \quad u_0 = \begin{bmatrix} 6 \\ 0 \end{bmatrix}.$$

32. Diagonalice A y calcule $S\Lambda^k S^{-1}$ con la finalidad de demostrar esta fórmula para A^k:

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{tiene} \quad A^k = \frac{1}{2}\begin{bmatrix} 3^k + 1 & 3^k - 1 \\ 3^k - 1 & 3^k + 1 \end{bmatrix}.$$

33. Diagonalice B y calcule $S\Lambda^k S^{-1}$ con la finalidad de demostrar esta fórmula para B^k:

$$B = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \quad \text{tiene} \quad B^k = \begin{bmatrix} 3^k & 3^k - 2^k \\ 0 & 2^k \end{bmatrix}.$$

Los problemas 34 a 44 son nuevas aplicaciones de $A = S\Lambda S^{-1}$.

34. Suponga que $A = S\Lambda S^{-1}$. Tome determinantes para demostrar que $A = \lambda_1\lambda_2 \cdots \lambda_n =$ producto de λs. Esta rápida demostración sólo funciona cuando A es _____.

35. La traza de S multiplicado por ΛS^{-1} es igual a la traza de ΛS^{-1} multiplicada por S. Así, la traza de una matriz diagonalizable A es igual a la traza de Λ, que es _____.

36. Si $A = S\Lambda S^{-1}$, diagonalice la matriz en bloque $B = \begin{bmatrix} A & 0 \\ 0 & 2A \end{bmatrix}$. Encuentre sus matrices valor característico y vector característico.

37. Considere todas las matrices A de 4 por 4 que son diagonalizadas por la misma matriz vector característico fija S. Demuestre que las As constituyen un subespacio (cA y $A_1 + A_2$ tienen esta misma S). ¿Cuál es este subespacio cuando $S = I$? ¿Cuál es su dimensión?

38. Suponga que $A^2 = A$. En el miembro izquierdo A multiplica cada columna de A. ¿Cuál de los cuatro subespacios contiene vectores característicos con $\lambda = 1$? ¿Qué subespacio contiene vectores característicos con $\lambda = 0$? Con base en las dimensiones de estos subespacios, A tiene un conjunto completo de vectores característicos independientes y es posible diagonalizarla.

39. Suponga que $Ax = \lambda x$. Si $\lambda = 0$, entonces x está en el espacio nulo. Si $\lambda \neq 0$, entonces x está en el espacio columna. Las dimensiones de estos subespacios son $(n - r) + r = n$. Entonces, ¿por qué ninguna matriz cuadrada tiene n vectores característicos linealmente independientes?

40. Sustituya $A = S\Lambda S^{-1}$ en el producto $(A - \lambda_1 I)(A - \lambda_2 I)\cdots(A - \lambda_n I)$, y explique por qué esto produce la *matriz cero*. Se está sustituyendo la matriz A por el número λ en el polinomio $p(\lambda) = \det(A - \lambda I)$. El ***teorema de Cayley-Hamilton*** establece que este producto siempre es $p(A) =$ *matriz cero*, incluso si A no es diagonalizable.

41. Pruebe el teorema de Cayley-Hamilton en la matriz de Fibonacci $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. El teorema pronostica que $A^2 - A - I = 0$, ya que $\det(A - \lambda I)$ es $\lambda^2 - \lambda - 1$.

42. Si $A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$, entonces $\det(A - \lambda I)$ es $(\lambda - a)(\lambda - d)$. Compruebe la afirmación del teorema de Cayley-Hamilton de que $(A - aI)(A - dI) =$ *matriz cero*.

43. Si $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ y $AB = BA$, demuestre que $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ también es diagonal. B tiene los mismos _____ característicos que A, pero diferentes _____ característicos. Estas matrices diagonales B constituyen un subespacio bidimensional del espacio de ma-

trices. $AB - BA = 0$ proporciona cuatro ecuaciones para las incógnitas **a**, **b**, **c**, **d**. Encuentre el rango de la matriz de 4 por 4.

44. Si A es de 5 por 5, entonces $AB - BA$ = matriz cero proporciona 25 ecuaciones para los 25 elementos de B. Demuestre que la matriz de 25 por 25 es singular, indicando una simple solución B diferente de cero.

45. Encuentre los valores característicos y los vectores característicos para las dos siguientes matrices A y A^∞ de Markov. Explique por qué A^{100} está próxima a A^∞:

$$A = \begin{bmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix} \quad \text{y} \quad A^\infty = \begin{bmatrix} 1/3 & 1/3 \\ 2/3 & 2/3 \end{bmatrix}.$$

5.3 ECUACIONES EN DIFERENCIAS Y POTENCIAS A^k

Las ecuaciones en diferencias $u_{k+1} = Au_k$ se mueven hacia adelante en un número finito de pasos finitos. Una ecuación diferencial asume una infinidad de pasos infinitesimales, aunque ambas teorías permanecen absolutamente en paralelo. Se trata de la misma analogía entre el discreto y el continuo que aparecen una y otra vez en matemáticas. Un buen ejemplo es el interés compuesto cuando el paso temporal se hace más corto.

Suponga que se invierten $1000 al 6% de interés. Si se compone una vez al año, el capital P se multiplica por 1.06. *Esta es una ecuación en diferencias* $P_{k+1} = AP_k = 1.06\, P_k$ *con un paso temporal de un año*. Al cabo de 5 años, la $P_0 = 1000$ original se ha multiplicado cinco veces:

$$\textbf{Anual} \qquad P_5 = (1.06)^5\, P_0 \quad \text{que es} \quad (1.06)^5\, 1000 = \$1338.$$

Ahora suponga que el paso temporal se reduce a un mes. La nueva ecuación en diferencias es $p_{k+1} = (1 + 0.06/12)p_k$. Luego de 5 años, o 60 meses, se tienen $11 más:

$$\textbf{Mensual} \qquad p_{60} = \left(1 + \frac{0.06}{12}\right)^{60} p_0 \quad \text{que es} \quad (1.005)^{60}\, 1000 = \$1349.$$

El paso siguiente es componer cada día, sobre 5(365) días. Esto es de poca ayuda:

$$\textbf{Composición diaria} \qquad \left(1 + \frac{0.06}{365}\right)^{5\cdot 365} 1000 = \$1349.83.$$

Finalmente, para mantener activos a sus empleados, los bancos ofrecen *composición continua*. El interés se suma a cada instante, y la ecuación en diferencias fracasa. Es de esperar que el tesorero no sepa cálculo (lo cual es todo sobre límites cuando $\Delta t \to 0$). El banco puede componer el interés N veces al año, de modo que $\Delta t = 1/N$:

$$\textbf{Continuamente} \qquad \left(1 + \frac{0.06}{N}\right)^{5N} 1000 \quad \to \quad e^{0.30}\, 1000 = \$1349.87.$$

O bien, el banco puede pasar a una ecuación diferencial: el límite de la ecuación en diferencias $p_{k+1} = (1 + 0.06\Delta t)p_k$. Al pasar p_k al miembro izquierdo y dividir entre Δt, se obtiene

$$\textbf{De discreto a continuo} \qquad \frac{p_{k+1} - p_k}{\Delta t} = 0.06 p_k \quad \text{tiende a} \quad \frac{dp}{dt} = 0.06p. \qquad (1)$$

La solución es $p(t) = e^{0.06t}p_0$. Al cabo de $t = 5$ años, lo anterior nuevamente suma $1349.87. El capital sigue siendo finito, aun cuando se componga a cada instante, y la mejoría sobre la composición diaria es de sólo cuatro centavos.

Números de Fibonacci

El objetivo principal de esta sección es resolver $u_{k+1} = Au_k$. Lo anterior lleva a A^k y a **potencias de matrices**. El segundo ejemplo es la famosa ***sucesión de Fibonacci***:

$$\textbf{Números de Fibonacci} \qquad 0, 1, 1, 2, 3, 5, 8, 13, \ldots$$

Este es el patrón: cada número es la suma de los dos números de Fibonacci previos Fs:

$$\textbf{Ecuación de Fibonacci} \qquad F_{k+2} = F_{k+1} + F_k. \qquad (2)$$

Ésta es la ecuación en diferencias. Aparece en una inmensa variedad de aplicaciones, de suyo merece todo un libro. Las hojas crecen siguiendo un patrón en espiral, y en el manzano o en el roble se encuentran cinco crecimientos por cada dos vueltas alrededor del tallo. El peral tiene ocho crecimientos por cada tres vueltas, y el sauce es 13:5. Parece que el campeón es un girasol cuyas semillas escogen una razón casi increíble de $F_{12}/F_{13} = 144/233$.*

¿Cómo es posible encontrar el milésimo término de Fibonacci sin empezar en $F_0 = 0$ y $F_1 = 1$, y trabajar hasta llegar a F_{1000}? El objetivo es resolver la ecuación en diferencias $F_{k+2} = F_{k+1} + F_k$. **Esta ecuación puede reducirse a una ecuación de un paso** $u_{k+1} = Au_k$. **Cada paso multiplica** $u_k = (F_{k+1}, F_k)$ **por una matriz** A:

$$\begin{array}{l} F_{k+2} = F_{k+1} + F_k \\ F_{k+1} = F_{k+1} \end{array} \quad \text{se convierte en} \quad u_{k+1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = Au_k. \qquad (3)$$

El sistema de un paso $u_{k+1} = Au_k$ es fácil de resolver. Empieza desde u_0. Luego de un paso se obtiene $u_1 = Au_0$. Luego, u_2 es Au_1, que es A^2u_0. *Cada paso conlleva una multiplicación por* A, y al cabo de k pasos hay k multiplicaciones:

> ***La solución de una ecuación en diferencias*** $u_{k+1} = Au_k$ **es** $u_k = A^k u_0$.

El verdadero problema consiste en encontrar alguna forma rápida para calcular el milésimo término de Fibonacci. La clave se encuentra en los valores característicos y en los vectores característicos:

> **5G** Si es posible diagonalizar A, $A = S\Lambda S^{-1}$, entonces A^k proviene de Λ^k:
>
> $$u_k = A^k u_0 = (S\Lambda S^{-1})(S\Lambda S^{-1})\cdots(S\Lambda S^{-1})u_0 = S\Lambda^k S^{-1}u_0. \qquad (4)$$
>
> Las columnas de S son los vectores característicos de A. Al escribir $S^{-1}u_0 = c$, la solución se convierte en
>
> $$u_k = S\Lambda^k c = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = c_1\lambda_1^k x_1 + \cdots + c_n\lambda_n^k x_n. \qquad (5)$$

Después de k pasos, u_k es una combinación de las n "soluciones puras" $\lambda^k x$.

*Para estas aplicaciones en botánica, consulte el libro *On Growth and Form* de D'Arcy Thompson (Cambridge University Press, 1942), o la hermosa obra de Peter Stevens, *Patterns in Nature* (Little, Brown, 1974). Cientos de otras propiedades de los F_n han sido publicadas en el *Fibonacci Quarterly*. Parece que fue Fibonacci quien llevó los números arábigos a Europa, alrededor de 1200 d.C.

Estas fórmulas constituyen dos métodos para encontrar la misma solución $u_k = S\Lambda^k S^{-1}u_0$.. La primera fórmula reconocía que A^k es idéntica con $S\Lambda^k S^{-1}$, y es posible detenerse ahí. Sin embargo, el segundo método conlleva la analogía con una ecuación diferencial: ***las soluciones puramente exponenciales*** $e^{\lambda_i t}x_i$ ***son ahora las potencias puras de*** $\lambda_i^k x_i$. Los vectores característicos x_i son amplificados por los valores característicos λ_i. Al combinar estas soluciones especiales para coincidir con u_0 —que es de donde proviene c— se recupera la solución correcta $u_k = S\Lambda^k S^{-1}u_0$.

En cualquier ejemplo específico como el de Fibonacci, el primer paso es encontrar los valores característicos:

$$A - \lambda I = \begin{bmatrix} 1-\lambda & 1 \\ 1 & -\lambda \end{bmatrix} \text{ tiene } \det(A - \lambda I) = \lambda^2 - \lambda - 1$$

$$\textbf{Dos valores característicos} \quad \lambda_1 = \frac{1+\sqrt{5}}{2} \quad \text{y} \quad \lambda_2 = \frac{1-\sqrt{5}}{2}.$$

El segundo renglón de $A - \lambda I$ es $(1, -\lambda)$. Para obtener $(A - \lambda I)x = 0$, el vector característico es $x = (\lambda, 1)$. Los primeros números de Fibonacci $F_0 = 0$ y $F_1 = 1$ van en u_0, y $S^{-1}u_0 = c$:

$$S^{-1}u_0 = \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ proporciona } c = \begin{bmatrix} 1/(\lambda_1 - \lambda_2) \\ -1/(\lambda_1 - \lambda_2) \end{bmatrix} = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Estas son las constantes en $u_k = c_1\lambda_1^k x_1 + c_2\lambda_2^k x_2$. Ambos vectores característicos x_1 y x_2 tienen por segunda componente a 1. Lo anterior deja $F_k = c_1\lambda_1^k + c_2\lambda_2^k$ en la segunda componente de u_k:

$$\textbf{Números de Fibonacci} \qquad F_k = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^k - \left(\frac{1-\sqrt{5}}{2}\right)^k\right].$$

Esta es la respuesta que se buscaba. Las fracciones y las raíces cuadradas se ven sorprendentes porque la regla de Fibonacci $F_{k+2} = F_{k+1} + F_k$ debe producir números enteros. De alguna manera esa fórmula para F_k debe proporcionar un entero. De hecho, como el segundo término $[(1 - \sqrt{5})/2]^k/\sqrt{5}$ siempre es menor que $\frac{1}{2}$, sólo debe mover el primer término al entero más próximo:

$$F_{1000} = \text{entero más próximo a } \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{1000}.$$

Este número es enorme, y F_{1001} es todavía más grande. Las fracciones se vuelven insignificantes, y el cociente F_{1001}/F_{1000} debe estar muy próximo a $(1 + \sqrt{5})/2 \approx 1.618$. Debido a que λ_2^k es insignificante en comparación con λ_1^k, el cociente F_{k+1}/F_k tiende a λ_1.

Esta es una ecuación en diferencias típica, que lleva a las potencias de $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. Implica $\sqrt{5}$ porque así lo hacen los valores característicos. Si se escoge una matriz con $\lambda_1 = 1$ y $\lambda_2 = 6$, es posible centrarse en la simpleza del cálculo, *después que se ha diagonalizado A*:

$$A = \begin{bmatrix} -4 & -5 \\ 10 & 11 \end{bmatrix} \text{ tiene } \lambda = 1 \text{ y } 6, \quad \text{con } x_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{y} \quad x_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$A^k = S\Lambda^k S^{-1} \quad \text{es} \quad \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} \mathbf{1}^k & 0 \\ 0 & \mathbf{6}^k \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 - 6^k & 1 - 6^k \\ -2 + 2\cdot 6^k & -1 + 2\cdot 6^k \end{bmatrix}.$$

Las potencias 6^k y 1^k aparecen en la última matriz A^k, mezcladas por los vectores característicos.

Para la ecuación en diferencias $u_{k+1} = Au_k$, se recalca la cuestión principal. Cada vector característico x produce una "solución pura" con potencias de λ:

$$\textit{Una solución es} \quad u_0 = x, \quad u_1 = \lambda x, \quad u_2 = \lambda^2 x, \ldots$$

Cuando el u_0 inicial es un vector característico x, esta es *la* solución: $u_k = \lambda^k x$. En general, u_0 no es un vector característico. Pero si u_0 es *una combinación* de vectores característicos, entonces la solución u_k es la misma combinación de estas soluciones especiales.

5H Si $u_0 = c_1x_1 + \cdots + c_nx_n$, entonces al cabo de k pasos $u_k = c_1\lambda_1^k x_1 + \cdots + c_n\lambda_n^k x_n$. Los cs se escogen para coincidir con el vector inicial u_0:

$$u_0 = \begin{bmatrix} & & \\ x_1 & \cdots & x_n \\ & & \end{bmatrix}\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = Sc \quad \text{y} \quad c = S^{-1}u_0. \tag{6}$$

Matrices de Markov

En el capítulo 1, hay un ejercicio sobre salir y entrar a California, que merece un análisis más detallado. Las reglas eran las siguientes:

Cada año, $\frac{1}{10}$ de la gente que vive fuera de California se cambia a este estado, y $\frac{2}{10}$ de la gente que vive en California sale del estado. Se empieza con y_0 gente fuera y z_0 dentro.

Al final del primer año las cantidades fuera y dentro son y_1 y z_1:

$$\textbf{Ecuación en diferencias} \quad \begin{aligned} y_1 &= 0.9y_0 + 0.2z_0 \\ z_1 &= 0.1y_0 + 0.8z_0 \end{aligned} \quad \text{o bien,} \quad \begin{bmatrix} y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}\begin{bmatrix} y_0 \\ z_0 \end{bmatrix}.$$

Este problema y su matriz poseen las dos propiedades esenciales de un ***proceso de Markov***:

1. El número total de gente permanece fijo. ***La suma de cada columna de la matriz de Markov es 1.*** No se gana ni pierde a nadie.
2. Los números de gente afuera y dentro jamás pueden volverse negativos: ***La matriz no tiene elementos negativos.*** Todas las potencias de A^k son no negativas.*

Esta ecuación en diferencias de Markov se resuelve, usando $u_k = S\Lambda^k S^{-1}u_0$. Luego se demuestra que la población tiende a un "estado estacionario". Primero es necesario diagonalizar a A:

$$A - \lambda I = \begin{bmatrix} 0.9 - \lambda & 0.2 \\ 0.1 & 0.8 - \lambda \end{bmatrix} \quad \text{tiene} \quad \det(A - \lambda I) = \lambda^2 - 1.7\lambda + 0.7$$

$$\boldsymbol{\lambda_1 = 1} \text{ y } \boldsymbol{\lambda_2} = 0.7: \qquad A = S\Lambda S^{-1} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{bmatrix}\begin{bmatrix} 1 & \\ & 0.7 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}.$$

*Además, la historia ha sido relegada por completo; cada nuevo u_{k+1} depende sólo del u_k actual. Tal vez incluso nuestras vidas son ejemplos de procesos de Markov, aunque espero que no lo sean.

Para encontrar A^k, y la distribución luego de k años, $S\Lambda S^{-1}$ se cambia por $S\Lambda^k S^{-1}$:

$$\begin{bmatrix} y_k \\ z_k \end{bmatrix} = A^k \begin{bmatrix} y_0 \\ z_0 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} 1^k & \\ & 0.7^k \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} y_0 \\ z_0 \end{bmatrix}$$

$$= (y_0 + z_0) \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} + (y_0 - 2z_0)(0.7)^k \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \end{bmatrix}.$$

Los dos términos son $c_1\lambda_1^k x_1 + c_2\lambda_2^k x_2$. El factor $\lambda_1^k = 1$ está escondido en el primer término. A largo plazo, el otro factor $(0.7)^k$ se vuelve extremadamente pequeño. ***La solución tiende a un estado limitante*** $u_\infty = (y_\infty, z_\infty)$:

$$\textbf{Estado estacionario} \qquad \begin{bmatrix} y_\infty \\ z_\infty \end{bmatrix} = (y_0 + z_0) \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}.$$

La población total sigue siendo $y_0 + z_0$, pero en el límite $\frac{2}{3}$ de esta población está fuera de California y $\frac{1}{3}$ está dentro. ¡Lo anterior es cierto sin importar cuál haya sido la distribución inicial! Si el año empieza con $\frac{2}{3}$ fuera y $\frac{1}{3}$ dentro, entonces termina de la misma forma:

$$\begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}, \quad \text{o bien}, \quad Au_\infty = u_\infty.$$

El estado estacionario es el vector característico de A correspondiente a $\lambda = 1$. Multiplicar por A, de un paso temporal al otro, deja sin cambio a u_∞.

La teoría de los procesos de Markov, se ilustra con este ejemplo de California:

5I Una matriz de Markov tiene a todos los $a_{ij} \geq 0$, donde la suma de cada columna es igual a 1.

a) $\lambda_1 = 1$ es un valor característico de A.
b) Su vector característico x_1 es no negativo, y es un estado estacionario porque $Ax_1 = x_1$.
c) Los otros valores característicos cumplen $|\lambda_i| \leq 1$.
d) Si A o cualquier otra potencia de A tiene a todos sus elementos *positivos*, estos otros $|\lambda_i|$ están abajo de 1.
La solución $A^k u_0$ se aproxima a un múltiplo de x_1, el cual es el estado estacionario u_∞.

Para encontrar el múltiplo idóneo de x_1, se usa el hecho de que la población total permanece igual. Si California empezó con todos sus 90 millones de gente fuera, entonces terminó con 60 millones fuera y 30 dentro. Termina de la misma forma si originalmente los 90 millones están dentro.

Se observa que muchos autores trasponen la matriz, de modo que la suma de sus *renglones* es 1.

Observación Esta descripción de un proceso de Markov fue determinista: las poblaciones se movieron en proporciones fijas. Pero si se considera un simple individuo, las fracciones que se mueven se convierten en *probabilidades*. Un individuo fuera de California entra con probabilidad $\frac{1}{10}$ a este estado. Si está dentro, la probabilidad de salir es $\frac{2}{10}$. El movimiento se vuelve un *proceso aleatorio*, y A se denomina ***matriz de transición***.

Las componentes de $u_k = A^k u_0$ especifican la probabilidad de que el individuo esté fuera o dentro del estado. Estas probabilidades nunca son negativas, y su suma es 1; es decir, todo mundo debe estar en algún sitio. Esto retrotrae las dos propiedades fundamentales de una matriz de Markov: La suma de cada columna es 1, y ningún elemento es negativo.

¿Por qué $\lambda = 1$ siempre es un valor característico? La suma de cada columna de $A - I$ es $1 - 1 = 0$. En consecuencia, la suma de los renglones de $A - I$ es el renglón cero, los renglones son linealmente dependientes, y $\det(A - I) = 0$.

Excepto por casos muy especiales, u_k tiende al vector característico correspondiente.* En la fórmula $u_k = c_1\lambda_1^k x_1 + \cdots + c_n\lambda_n^k x_n$ ningún valor característico puede ser mayor que 1. (En caso contrario se inflarían las probabilidades u_k). Si todos los demás valores característicos son estrictamente menores que $\lambda_1 = 1$, entonces el primer término en la fórmula es dominante. Los otros λ_i^k se van a cero, y $u_k \to c_1x_1 = u_\infty$ = estado estacionario.

Este es un ejemplo de uno de los temas torales de este capítulo: Dada información sobre A, encontrar información sobre sus valores característicos. Aquí se encontró $\lambda_{\text{máx}} = 1$.

Estabilidad de $u_{k+1} = Au_k$

Hay una diferencia evidente entre los números de Fibonacci y los procesos de Markov. Los números F_k se vuelven cada vez más grandes, mientras la definición de "probabilidad" está entre 0 y 1. La ecuación de Fibonacci es *inestable*. También lo es la ecuación del interés compuesto $P_{k+1} = 1.06P_k$; el capital crece por siempre. Si las probabilidades de Markov decrecen a cero, esa ecuación sería estable; pero no lo hacen, ya que en cada etapa su suma debe ser 1. En consecuencia, un proceso de Markov es *neutralmente estable*.

Se desea estudiar el comportamiento de $u_{k+1} = Au_k$ cuando $k \to \infty$. Suponiendo que A puede diagonalizarse, u_k es una combinación de soluciones puras:

$$\textbf{Solución en el instante } k \qquad u_k = S\Lambda^k S^{-1}u_0 = c_1\lambda_1^k x_1 + \cdots + c_n\lambda_n^k x_n.$$

El crecimiento de u_k está regido por los λ_i^k. ***La estabilidad depende de los valores característicos:***

5J La ecuación en diferencias $u_{k+1} = Au_k$ es

estable si todos los valores característicos cumplen $|\lambda_i| < 1$.
neutralmente estable si algunos $|\lambda_i| = 1$ y los demás $|\lambda_i| < 1$; e
inestable si por lo menos un valor característico cumple $|\lambda_i| > 1$.

En el caso estable, las potencias A^k tienden a cero, así como $u_k = A^k u_0$.

Ejemplo 1 La matriz A es ciertamente estable:

$$A = \begin{bmatrix} 0 & 4 \\ 0 & \frac{1}{2} \end{bmatrix} \text{ tiene valores característicos } 0 \text{ y } \tfrac{1}{2}.$$

Los λs están en la diagonal principal porque A es triangular. Empezando con cualquier u_0, y siguiendo la regla $u_{k+1} = Au_k$, la solución debe terminar por tender a cero:

$$u_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad u_1 = \begin{bmatrix} 4 \\ \frac{1}{2} \end{bmatrix}, \quad u_2 = \begin{bmatrix} 2 \\ \frac{1}{4} \end{bmatrix}, \quad u_3 = \begin{bmatrix} 1 \\ \frac{1}{8} \end{bmatrix}, \quad u_4 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{16} \end{bmatrix}, \quad \cdots$$

*Si todos los que están afuera se mueven hacia dentro y todos los que están adentro se mueven hacia fuera, entonces las poblaciones se invierten cada año y no hay estado estacionario. La matriz de transición es $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ y -1 es un valor característico, así como $+1$, lo cual no puede ocurrir si todos los $a_{ij} > 0$.

El mayor valor característico $\lambda = \frac{1}{2}$ rige la disminución; después del primer paso, cada u_k es $\frac{1}{2}u_{k-1}$. El verdadero efecto del primer paso, es separar u_0 en los dos vectores característicos de A:

$$u_0 = \begin{bmatrix} 8 \\ 1 \end{bmatrix} + \begin{bmatrix} -8 \\ 0 \end{bmatrix} \quad \text{y entonces} \quad u_k = \left(\frac{1}{2}\right)^k \begin{bmatrix} 8 \\ 1 \end{bmatrix} + (0)^k \begin{bmatrix} -8 \\ 0 \end{bmatrix}.$$

Matrices positivas y aplicaciones a la economía

Al desarrollar las ideas de Markov es posible encontrar una pequeña mina de oro (*completamente opcional*) de aplicaciones matriciales a la economía.

Ejemplo 2 *Matriz entrada-salida de Leontief*

Este es uno de los primeros grandes éxitos de la economía matemática. Para ilustrarlo, se construye una *matriz consumo,* en la que a_{ij} proporciona la cantidad del producto j que se requiere para obtener una unidad del producto i:

$$A = \begin{bmatrix} 0.4 & 0 & 0.1 \\ 0 & 0.1 & 0.8 \\ 0.5 & 0.7 & 0.1 \end{bmatrix}. \quad \begin{matrix} \text{(acero)} \\ \text{(alimentos)} \\ \text{(trabajo)} \end{matrix}$$

La primera pregunta es: ¿Es posible producir y_1 unidades de acero, y_2 unidades de alimentos, y y_3 unidades de trabajo? Es necesario empezar con grandes cantidades p_1, p_2, p_3, porque una porción se consume en la producción misma. La cantidad que se consume es Ap, y deja una producción neta de $p - Ap$.

Problema *Encontrar un vector p tal que* $p - Ap = y$, *o bien,* $p = (I - A)^{-1}y$.

Aparentemente sólo se está preguntando si $I - A$ es invertible. Sin embargo, hay un giro negativo del problema. La demanda y la producción, y y p, son no negativas. Debido a que p es $(I - A)^{-1}y$, la verdadera pregunta es sobre la matriz que multiplica a y:

¿Cuándo ocurre que $(I - A)^{-1}$ es una matriz no negativa?

Aproximadamente, A no puede ser demasiado grande. Si la producción consume demasiado, nada queda como salida. La clave radica en el mayor valor característico λ_1 de A, que debe ser menor que 1:

si $\lambda_1 > 1$, $(I - A)^{-1}$ fracasa en ser no negativa.
si $\lambda_1 = 1$, $(I - A)^{-1}$ fracasa en existir.
si $\lambda_1 < 1$, $(I - A)^{-1}$ es una suma convergente de matrices no negativas:

Serie geométrica $$(I - A)^{-1} = I + A + A^2 + A^3 + \cdots \qquad (7)$$

En el ejemplo de 3 por 3 se tenía $\lambda_1 = 0.9$, y la salida excedía a la entrada. La producción puede continuar.

Es fácil demostrar lo anterior, una vez que se conoce el hecho principal sobre una matriz no negativa como A: ***No sólo el mayor valor característico λ_1 es positivo, sino que también lo es el vector característico x_1.*** Así $(I - A)^{-1}$ tiene el mismo vector característico, con valor característico $1/(1 - \lambda_1)$.

Si λ_1 excede a 1, ese último número es negativo. La matriz $(I - A)^{-1}$ tomará el vector positivo x_1 como si fuese un vector negativo $x_1/(1 - \lambda_1)$. En ese caso $(I - A)^{-1}$ definitivamente no es no negativa. Si $\lambda_1 = 1$, entonces $I - A$ es singular. El caso productivo es $\lambda_1 < 1$, cuando las potencias de A se hacen cero (estabilidad) y la serie infinita $I + A + A^2 + \ldots$ converge.

Al multiplicar esta serie por $I - A$ se queda con la matriz identidad; todas las potencias superiores se cancelan, de modo que $(I - A)^{-1}$ es una suma de matrices no negativas. A continuación se proporcionan dos ejemplos:

$$A = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix} \text{ tiene } \lambda_1 = 2 \text{ y se pierde la economía}$$

$$A = \begin{bmatrix} 0.5 & 2 \\ 0 & 0.5 \end{bmatrix} \text{ tiene } \lambda_1 = \frac{1}{2} \text{ y es imposible producir algo.}$$

En estos casos, las matrices $(I - A)^{-1}$ son $-\frac{1}{3}\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ y $\begin{bmatrix} 2 & 8 \\ 0 & 2 \end{bmatrix}$.

La inspiración de Leontief fue encontrar un modelo en que se utilizan datos genuinos de la economía real. La tabla para 1958 contenía 83 industrias en Estados Unidos, con una "tabla de transacciones" de consumo y producción para cada una. La teoría también llega más allá de $(I - A)^{-1}$, para decidir precios naturales y cuestiones de optimización. Normalmente, el trabajo está limitado en suministro y debe minimizarse. Y, por supuesto, la economía no siempre es lineal.

Ejemplo 3 ***Los precios en un modelo entrada-salida*** *cerrado*

El modelo se denomina "cerrado" cuando todo lo que se produce es consumido. Nada sale del sistema. En ese caso, A regresa a ser una *matriz de Markov*. ***La suma de las columnas es 1.*** Podría hablarse sobre el *valor* del acero, y de los alimentos, y del trabajo, en vez del número de unidades. El vector p representa precios, en vez de niveles de producción.

Suponga que p_0 es un vector de precios. Entonces Ap_0 multiplica los precios por cantidades con la finalidad de proporcionar el valor de cada producto. Este es un nuevo conjunto de precios que el sistema utiliza para el siguiente conjunto de valores A^2p_0. La pregunta es si los precios tienden al equilibrio. ¿Existen precios tales que $p = Ap$, y el sistema conduce a ellos?

p se reconoce como el vector característico (no negativo) de la matriz A de Markov, con $\lambda = 1$. Este es el estado estacionario p_∞, al que se tiende desde cualquier punto inicial p_0. Al repetir una transacción una y otra vez, los precios tienden al equilibrio.

El teorema de "Perron-Frobenius" proporciona las propiedades clave de una ***matriz positiva***, que no debe confundirse con una matriz *positiva definida*, que es simétrica y todos sus valores característicos son positivos. Aquí todos los elementos a_{ij} son positivos.

5K Si A es una matriz positiva, también lo es su mayor valor característico: $\lambda_1 >$ todos los demás $|\lambda_i|$. Cada componente del vector característico x_1 correspondiente, también es positivo.

Demostración Suponga $A > 0$. La idea clave es considerar todos los números t tales que $Ax \geq tx$ para algún vector no negativo x (que no sea $x = 0$). La desigualdad $Ax \geq tx$ se deja para contar con muchos candidatos positivos t. Para el mayor valor $t_{\text{máx}}$ (que se alcanza), se demostrará que ***la desigualdad se cumple***: $Ax = t_{\text{máx}}x$.

En caso contrario, si $Ax \geq t_{\text{máx}}x$ no es una igualdad, se multiplica por A. Debido a que A es positiva, lo anterior produce una desigualdad estricta $A^2x > t_{\text{máx}}Ax$. En consecuencia, el vector positivo $y = Ax$ cumple $Ay > t_{\text{máx}}y$, y $t_{\text{máx}}$ hubiera podido ser más grande. Esta contradicción obliga a la igualdad $Ax = t_{\text{máx}}x$, y se tiene un valor característico. Su vector característico x es positivo (no solamente no negativo) debido a que en el miembro izquierdo de esa desigualdad, es seguro que Ax es positivo.

Para ver que ningún valor característico puede ser mayor que $t_{máx}$, suponga que $Az = \lambda z$. Debido a que tanto λ como z pueden implicar números negativos o complejos, se toman valores absolutos: $|\lambda||z| = |Az| \leq A|z|$ por la "desigualdad del triángulo". Este $|z|$ es un vector no negativo, de modo que $|\lambda|$ es uno de los candidatos posibles t. En consecuencia $|\lambda|$ no puede exceder a λ_1, que era $t_{máx}$. ■

Ejemplo 4 ***Modelo de von Neumann, de una economía en expansión***

Se regresará a la matriz A de 3 por 3 que proporcionó el consumo de combustible, alimentos, y trabajo. Si las salidas son s_1, f_1, ℓ_1, entonces las entradas requeridas son

$$u_0 = \begin{bmatrix} 0.4 & & 0.1 \\ 0 & 0.1 & 0.8 \\ 0.5 & 0.7 & 0.1 \end{bmatrix} \begin{bmatrix} s_1 \\ f_1 \\ \ell_1 \end{bmatrix} = Au_1.$$

En economía, ¡la ecuación en diferencias es hacia atrás! En vez de $u_1 = Au_0$, se tiene $u_0 = Au_1$. Si A es pequeña (como es el caso), entonces la producción no consume todo, y la economía puede crecer. Los valores característicos de A^{-1} rigen este crecimiento. Pero otra vez hay un giro no negativo, ya que el acero, los alimentos, y el trabajo no se presentan en cantidades negativas. Von Neumann se preguntó por la razón máxima t a la que puede crecer la economía *sin dejar de ser no negativa*, lo cual significa que $u_1 \geq tu_0 \geq 0$.

Así, el problema requiere $u_1 \geq tAu_1$. Es como el teorema de Perron-Frobenius, con A en el otro lado. Como antes, la igualdad se cumple cuando t alcanza $t_{máx}$; que es el valor característico asociado con el vector característico positivo de A^{-1}. En este caso, el factor de expansión es $\frac{10}{9}$:

$$x = \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix} \quad \text{y} \quad Ax = \begin{bmatrix} 0.4 & 0 & 0.1 \\ 0 & 0.1 & 0.8 \\ 0.5 & 0.7 & 0.1 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix} = \begin{bmatrix} 0.9 \\ 4.5 \\ 4.5 \end{bmatrix} = \frac{9}{10}x.$$

Con acero−alimentos−trabajo, en la razón 1−5−5, la economía crece tan rápido como es posible: ***La razón de crecimiento máxima es*** $\mathbf{1/\lambda_1}$.

Conjunto de problemas 5.3

1. Demuestre que cada tercer número de Fibonacci en 0, 1, 1, 2, 3, . . . es par.

2. Bernadelli estudió un escarabajo que "sólo vive tres años, y se reproduce en su tercer año". Estos escarabajos sobreviven el tercer año con probabilidad $\frac{1}{2}$, y el segundo con probabilidad $\frac{1}{3}$, y luego producen seis hembras al final de su vida:

$$\textbf{Matriz escarabajo} \qquad A = \begin{bmatrix} 0 & 0 & 6 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \end{bmatrix}.$$

Demuestre que $A^3 = I$, y siga la distribución de 3000 escarabajos durante seis años.

3. Para la matriz de Fibonacci $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, calcule A^2, A^3, A^4. Luego, use el texto y una calculadora, para encontrar F_{20}.

4. Suponga que cada número de "Gibonacci" G_{k+2} es el promedio de los dos números previos, G_{k+1} y G_k. Entonces $G_{k+2} = \frac{1}{2}(G_{k+1} + G_k)$:

$$\begin{aligned} G_{k+2} &= \tfrac{1}{2}G_{k+1} + \tfrac{1}{2}G_k \\ G_{k+1} &= G_{k+1} \end{aligned} \qquad \text{es} \qquad \begin{bmatrix} G_{k+2} \\ G_{k+1} \end{bmatrix} = \begin{bmatrix} & A & \end{bmatrix} \begin{bmatrix} G_{k+1} \\ G_k \end{bmatrix}.$$

a) Encuentre los valores característicos y los vectores característicos de A.
b) Encuentre el límite cuando $n \to \infty$ de las matrices $A^n = S\Lambda^n S^{-1}$.
c) Si $G_0 = 0$ y $G_1 = 1$, demuestre que los números de Gibonacci tienden a $\frac{2}{3}$.

5. Diagonalice la matriz de Fibonacci completando S^{-1}:

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} & \\ & \end{bmatrix}.$$

Realice la multiplicación $S\Lambda^k S^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ para encontrar su segunda componente. Este es el k-ésimo número de Fibonacci $F_k = (\lambda_1^k - \lambda_2^k)/(\lambda_1 - \lambda_2)$.

6. Los números λ_1^k y λ_2^k cumplen la regla de Fibonacci $F_{k+2} = F_{k+1} + F_k$:

$$\lambda_1^{k+2} = \lambda_1^{k+1} + \lambda_1^k \quad \text{y} \quad \lambda_2^{k+2} = \lambda_2^{k+1} + \lambda_2^k.$$

Demuestre lo anterior, usando la ecuación original para los λs (multiplíquela por λ^k). Entonces, cualquier combinación de λ_1^k y λ_2^k cumple la regla. La combinación $F_k = (\lambda_1^k - \lambda_2^k)/(\lambda_1 - \lambda_2)$ proporciona el inicio correcto de $F_0 = 0$ y $F_1 = 1$.

7. Lucas empezó con $L_0 = 2$ y $L_1 = 1$. La regla $L_{k+2} = L_{k+1} + L_k$ es la misma, de modo que A sigue siendo una matriz de Fibonacci. Sume sus vectores característicos $x_1 + x_2$:

$$\begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} + \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(1 + \sqrt{5}) \\ 1 \end{bmatrix} + \begin{bmatrix} \frac{1}{2}(1 - \sqrt{5}) \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} L_1 \\ L_0 \end{bmatrix}.$$

Al multiplicar por A^k, la segunda componente es $L_k = \lambda_1^k + \lambda_2^k$. Calcule el número de Lucas L_{10}, lentamente aplicando $L_{k+2} = L_{k+1} + L_k$, y calcule aproximadamente con λ_1^{10}.

8. Suponga que hay una epidemia en la cual, cada mes la mitad de los sanos enferman y la cuarta parte de los enfermos fallecen. Encuentre el estado estacionario para el proceso de Markov, correspondiente

$$\begin{bmatrix} d_{k+1} \\ s_{k+1} \\ w_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{4} & 0 \\ 0 & \frac{3}{4} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} d_k \\ s_k \\ w_k \end{bmatrix}.$$

9. Escriba la matriz de transición de 3 por 3 para un curso de química que se enseña en dos secciones, si cada semana $\frac{1}{4}$ de los de la sección A y $\frac{1}{3}$ de los de la sección B abandonan el curso, y $\frac{1}{6}$ de cada sección se cambia a la otra sección.

10. Encuentre los valores limitantes de y_k y z_k $(k \to \infty)$ si

$$\begin{aligned} y_{k+1} &= 0.8y_k + 0.3z_k & y_0 &= 0 \\ z_{k+1} &= 0.2y_k + 0.7z_k & z_0 &= 5. \end{aligned}$$

También, encuentre fórmulas para y_k y z_k, a partir de $A^k = S\Lambda^k S^{-1}$.

11. *a*) A partir del hecho de que columna 1 + columna 2 = 2(columna 3), de modo que las columnas son linealmente independientes, encuentre un valor característico y un vector característico de A:

$$A = \begin{bmatrix} 0.2 & 0.4 & 0.3 \\ 0.4 & 0.2 & 0.3 \\ 0.4 & 0.4 & 0.4 \end{bmatrix}.$$

b) Encuentre los otros valores característicos de A (es una matriz de Markov).
c) Si $u_0 = (0, 10, 0)$, encuentre el límite de $A^k u_0$ cuando $k \to \infty$.

12. Suponga que hay tres centros principales para los camiones Muévalo-Usted-Mismo. Cada mes, la mitad de los camiones que están en Boston y en Los Ángeles, van a Chicago, la otra mitad se queda donde están, y los camiones en Chicago se dividen equitativamente entre Boston y Los Ángeles. Escriba la matriz de transición A de 3 por 3, y encuentre el estado estacionario u_∞ correspondiente al valor característico $\lambda = 1$.

13. *a*) ¿En qué rango de a y b está la siguiente ecuación de un proceso de Markov?

$$u_{k+1} = Au_k = \begin{bmatrix} a & b \\ 1-a & 1-b \end{bmatrix} u_k, \qquad u_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

b) Calcule $u_k = S\Lambda^k S^{-1} u_0$ para cualesquiera a y b.
c) ¿En qué condición sobre a y b, u_k tiende a un límite finito cuando $k \to \infty$ y cuál es el límite? ¿A debe ser una matriz de Markov?

14. Empresas multinacionales en América, Asia y Europa poseen bienes por \$4 trillones. Al principio, \$2 trillones están en América y \$2 trillones están en Europa. Cada año, $\frac{1}{2}$ del dinero en América se queda en ese continente, y $\frac{1}{4}$ se va a Asia y Europa. Para Asia y Europa, $\frac{1}{2}$ permanece ahí y $\frac{1}{2}$ se envía a América.
a) Encuentre la matriz que proporciona

$$\begin{bmatrix} \text{América} \\ \text{Asia} \\ \text{Europa} \end{bmatrix}_{\text{año } k+1} = A \begin{bmatrix} \text{América} \\ \text{Asia} \\ \text{Europa} \end{bmatrix}_{\text{año } k}.$$

b) Encuentre los valores característicos y los vectores característicos de A.
c) Encuentre la distribución limitante de los \$4 trillones, cuando se acaba el mundo.
d) Encuentre la distribución de los \$4 trillones en el año k.

15. Si A es una matriz de Markov, demuestre que la suma de las componentes de Ax es igual a la suma de las componentes de x. Deduzca que si $Ax = \lambda x$ con $\lambda \neq 1$, entonces la suma de las componentes del vector característico es igual a cero.

16. La solución de $du/dt = Au = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ (valores característicos i y $-i$) gira alrededor de un círculo: $u = (\cos t, \operatorname{sen} t)$. Suponga que du/dt se aproxima por diferencias hacia adelante, hacia atrás, y centradas **F, B, C:**

(**F**) $u_{n+1} - u_n = Au_n$ o $u_{n+1} = (I + A)u_n$ (este es el método de Euler).
(**B**) $u_{n+1} - u_n = Au_{n+1}$ o $u_{n+1} = (I - A)^{-1}u_n$ (hacia atrás Euler).
(**C**) $u_{n+1} - u_n = \frac{1}{2}A(u_{n+1} + u_n)$ o $u_{n+1} = \left(I - \frac{1}{2}A\right)^{-1}\left(I + \frac{1}{2}A\right)u_n$.

Encuentre los valores característicos de $I + A$, $(I - A)^{-1}$, e $\left(I - \frac{1}{2}A\right)^{-1}\left(I + \frac{1}{2}A\right)$. ¿Para cuál ecuación en diferencias la solución u_n permanece en un círculo?

17. ¿Qué valores de α producen inestabilidad en $v_{n+1} = \alpha(v_n + w_n)$, $w_{n+1} = \alpha(v_n + w_n)$?

18. Encuentre los a, b, c más grandes, para los cuales las siguientes matrices son estables o neutralmente estables:

$$\begin{bmatrix} a & -0.8 \\ 0.8 & 0.2 \end{bmatrix}, \quad \begin{bmatrix} b & 0.8 \\ 0 & 0.2 \end{bmatrix}, \quad \begin{bmatrix} c & 0.8 \\ 0.2 & c \end{bmatrix}.$$

19. Multiplique término a término para comprobar que $(I - A)(I + A + A^2 + \cdots) = I$. Esta serie representa $(I - A)^{-1}$. Es no negativa cuando A es no negativa, en el supuesto de que tenga una suma finita; la condición para esto es $\lambda_{\text{máx}} < 1$. Sume la serie infinita, y confirme que es igual a $(I - A)^{-1}$, para la matriz de consumo

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{que tiene } \lambda_{\text{máx}} = 0.$$

20. Para $A = \begin{bmatrix} 0 & 0.2 \\ 0 & 0.5 \end{bmatrix}$, encuentre las potencias A^k (incluyendo A^0), y demuestre explícitamente que su suma coincide con $(I - A)^{-1}$.

21. Explique por matemáticas o economía por qué al incrementar la "matriz de consumo" A debe incrementarse $t_{\text{máx}} = \lambda_1$ (y aminorar el desarrollo).

22. ¿Cuáles son los límites cuando $k \to \infty$ (los estados estacionarios) de lo siguiente?

$$\begin{bmatrix} 0.4 & 0.2 \\ 0.6 & 0.8 \end{bmatrix}^k \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0.4 & 0.2 \\ 0.6 & 0.8 \end{bmatrix}^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0.4 & 0.2 \\ 0.6 & 0.8 \end{bmatrix}^k.$$

Los problemas 23 a 29 son sobre $A = S\Lambda S^{-1}$ y $A^k = S\Lambda^k S^{-1}$.

23. Diagonalice A y calcule $S\Lambda^k S^{-1}$ con la finalidad de demostrar esta fórmula para A^k:

$$A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} \quad \text{tiene} \quad A^k = \frac{1}{2}\begin{bmatrix} 5^k + 1 & 5^k - 1 \\ 5^k - 1 & 5^k + 1 \end{bmatrix}.$$

24. Diagonalice B y calcule $S\Lambda^k S^{-1}$ con la finalidad de demostrar esta fórmula para B^k:

$$B = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \quad \text{tiene} \quad B^k = \begin{bmatrix} 3^k & 3^k - 2^k \\ 0 & 2^k \end{bmatrix}.$$

25. Los valores característicos de A son 1 y 9, y los valores característicos de B son -1 y 9:

$$A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 4 & 5 \\ 5 & 4 \end{bmatrix}.$$

Encuentre una matriz raíz cuadrada de A, a partir de $R = S\sqrt{\Lambda}\,S^{-1}$. ¿Por qué para B no existe ninguna matriz raíz cuadrada real?

26. Si A y B tienen los mismos λs con el mismo conjunto completo de vectores característicos independientes, su factorización en ______ es la misma. Así, $A = B$.

27. Suponga que A y B tienen el mismo conjunto completo de vectores característicos, de modo que $A = S\Lambda_1 S^{-1}$ y $B = S\Lambda_2 S^{-1}$. Demuestre que $AB = BA$.

28. *a*) ¿Cuándo ocurre, que los vectores característicos para $\lambda = 0$ generan el espacio nulo $N(A)$?
b) ¿Cuándo ocurre, que todos vectores característicos para $\lambda \neq 0$ generan el espacio columna $C(A)$?

29. Las potencias de A^k tienden a cero si todos los $|\lambda_i| < 1$, y se inflan si cualquier $|\lambda_i| > 1$. Peter Lax proporciona cuatro ejemplos sorprendentes en su libro *Linear Algebra.*

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & 2 \\ -5 & -3 \end{bmatrix} \qquad C = \begin{bmatrix} 5 & 7 \\ -3 & -4 \end{bmatrix} \qquad D = \begin{bmatrix} 5 & 6.9 \\ -3 & -4 \end{bmatrix}$$

$$\|A^{1024}\| > 10^{700} \qquad B^{1024} = I \qquad C^{1024} = -C \qquad \|D^{1024}\| < 10^{-78}$$

Encuentre los valores característicos $\lambda = e^{i\theta}$ de B y C, para demostrar que $B^4 = I$ y que $C^3 = -I$.

5.4 ECUACIONES DIFERENCIALES Y e^{At}

Siempre que se encuentra un sistema de ecuaciones, más que una sola ecuación, la teoría de matrices tiene un papel que desempeñar. Para ecuaciones en diferencias, la solución $u_k = A^k u_0$ dependía de las potencias de A. Para ecuaciones diferenciales, la solución $u(t) = e^{At}u(0)$ depende de la ***exponencial*** de A. Para definir esta exponencial, y comprenderla, de inmediato se presenta un ejemplo:

$$\textbf{Ecuación diferencial} \qquad \frac{du}{dt} = Au = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} u. \qquad (1)$$

El primer paso siempre consiste en encontrar los valores característicos (-1 y -3), así como los vectores característicos:

$$A\begin{bmatrix} 1 \\ 1 \end{bmatrix} = (-1)\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{y} \quad A\begin{bmatrix} 1 \\ -1 \end{bmatrix} = (-3)\begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Luego, varios métodos llevan a $u(t)$. Quizá lo mejor sea hacer coincidir la solución general con el vector inicial $u(0)$ en $t = 0$.

La solución general es una combinación de soluciones exponenciales puras. Estas son soluciones de la forma especial $ce^{\lambda t}x$, donde λ es un valor característico de A, y x es su vector característico. Estas soluciones puras satisfacen la ecuación diferencial, ya que $d/dt(ce^{\lambda t}x) = A(ce^{\lambda t}x)$. (Fueron la introducción a los valores característicos al inicio del capítulo). En este ejemplo de 2 por 2, hay dos exponenciales puras a combinar:

$$\textbf{Solución} \qquad u(t) = c_1 e^{\lambda_1 t}x_1 + c_2 e^{\lambda_2 t}x_2 \quad \text{o bien,} \quad u = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} e^{-t} & \\ & e^{-3t} \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}. \qquad (2)$$

En el instante cero, cuando las exponenciales son $e^0 = 1$, $u(0)$ determina c_1 y c_2:

$$\textbf{Condición inicial} \qquad u(0) = c_1x_1 + c_2x_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = Sc.$$

Se reconoce S, la matriz de vectores característicos. Las constantes $c = S^{-1}u(0)$ son las mismas que para las ecuaciones en diferencias. Al sustituirlas de nuevo en la ecuación (2), se encuentra que la solución es

$$u(t) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} e^{-t} & \\ & e^{-3t} \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = S\begin{bmatrix} e^{-t} & \\ & e^{-3t} \end{bmatrix}S^{-1}u(0). \qquad (3)$$

He aquí la fórmula fundamental de esta sección: $Se^{\Lambda t}S^{-1}u(0)$ resuelve la ecuación diferencial, justo como $S\Lambda^k S^{-1}u_0$ resolvía la ecuación en diferencias:

$$u(t) = Se^{\Lambda t}S^{-1}u(0) \quad \text{con} \quad \Lambda = \begin{bmatrix} -1 & \\ & -3 \end{bmatrix} \quad \text{y} \quad e^{\Lambda t} = \begin{bmatrix} e^{-t} & \\ & e^{-3t} \end{bmatrix}. \qquad (4)$$

Hay dos cosas más que hacer con este ejemplo. Una es completar el procedimiento matemático, proporcionando una definición directa de la ***exponencial de una matriz.*** La otra es proporcionar una interpretación física de la ecuación y su solución. Se trata del tipo de ecuación diferencial que tiene aplicaciones de utilidad.

La exponencial de una matriz diagonal Λ es fácil: $e^{\Lambda t}$ tiene justo los n números $e^{\lambda t}$ en la diagonal. Para una matriz general A, la idea natural es imitar la serie de potencias $e^x = 1 + x + x^2/2! + x^3/3! + \cdots$. Si se sustituye x por At y 1 por I, esta suma es una matriz de n por n:

$$\textbf{Matriz exponencial} \qquad e^{At} = I + At + \frac{(At)^2}{2!} + \frac{(At)^3}{3!} + \cdots. \qquad (5)$$

La serie siempre es convergente, y su suma e^{At} tiene las propiedades idóneas:

$$(e^{As})(e^{At}) = e^{A(s+t)}, \qquad (e^{At})(e^{-At}) = I, \qquad \text{y} \qquad \frac{d}{dt}(e^{At}) = Ae^{At}. \qquad (6)$$

A partir de la última expresión, $u(t) = e^{At}u(0)$ resuelve la ecuación diferencial. Esta solución debe ser la misma que la forma $Se^{\Lambda t}S^{-1}u(0)$ empleada para computación. Para demostrar directamente que estas soluciones coinciden, recuerde que cada potencia $(S\Lambda S^{-1})^k$ se traduce en $A^k = S\Lambda^k S^{-1}$ (debido a que S^{-1} cancela a S). Toda la exponencial es diagonalizada por S:

$$e^{At} = I + S\Lambda S^{-1}t + \frac{S\Lambda^2S^{-1}t^2}{2!} + \frac{S\Lambda^3S^{-1}t^3}{3!} + \cdots$$

$$= S\left(I + \Lambda t + \frac{(\Lambda t)^2}{2!} + \frac{(\Lambda t)^3}{3!} + \cdots\right)S^{-1} = Se^{\Lambda t}S^{-1}.$$

Ejemplo 1 En la ecuación (1), la exponencial de $A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$ tiene $\Lambda = \begin{bmatrix} -1 & \\ & -3 \end{bmatrix}$:

$$e^{At} = Se^{\Lambda t}S^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} e^{-t} & \\ & e^{-3t} \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} = \frac{1}{2}\begin{bmatrix} e^{-t} + e^{-3t} & e^{-t} - e^{-3t} \\ e^{-t} - e^{-3t} & e^{-t} + e^{-3t} \end{bmatrix}.$$

En $t = 0$, se obtiene $e^0 = I$. La serie infinita e^{At} proporciona la respuesta para toda t, aunque es difícil calcular una serie. La forma $Se^{\Lambda t}S^{-1}$ proporciona la misma respuesta cuando es posible diagonalizar a A: se requieren n vectores característicos independientes en S. Esta forma más sencilla, lleva a una *combinación de n exponenciales* $e^{\lambda t}x$, que es la mejor solución de todas:

5L Si es posible diagonalizar a A, $A = S\Lambda S^{-1}$, entonces $du/dt = Au$ tiene la solución

$$u(t) = e^{At}u(0) = Se^{\Lambda t}S^{-1}u(0). \qquad (7)$$

Las columnas de S son los vectores característicos $x_1, \ldots, x_n$ de A. Al multiplicar se obtiene

$$u(t) = \begin{bmatrix} & & \\ x_1 & \cdots & x_n \\ & & \end{bmatrix}\begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix}S^{-1}u(0)$$

$$= c_1e^{\lambda_1 t}x_1 + \cdots + c_ne^{\lambda_n t}x_n = \text{combinación de } e^{\lambda t}x. \qquad (8)$$

Las constantes c_i que cumplen las condiciones iniciales $u(0)$ son $c = S^{-1}u(0)$.

Lo anterior proporciona una analogía completa con las ecuaciones en diferencias y $S\Lambda S^{-1}u_0$. En ambos casos se supuso que era posible diagonalizar a A, ya que en caso contrario posee menos de n vectores característicos y no se han encontrado suficientes soluciones especiales. Las soluciones que faltan existen, aunque son más complicadas que las exponenciales puras $e^{\lambda t}x$. Implican "vectores característicos generalizados" y factores como $te^{\lambda t}$. (Para calcular este caso defectuoso, puede usarse la forma de Jordan que se muestra en el apéndice B, y encontrar e^{Jt}). **La fórmula $u(t) = e^{At}u(0)$ sigue siendo completamente correcta.**

La matriz e^{At} ***nunca es singular***. Una demostración consiste en considerar sus valores característicos; si λ es un valor característico de A, entonces $e^{\lambda t}$ es el valor característico correspondiente de e^{At}, y $e^{\lambda t}$ jamás puede ser cero. Otro método es calcular el determinante de la exponencial:

$$\det e^{At} = e^{\lambda_1 t}e^{\lambda_2 t}\cdots e^{\lambda_n t} = e^{\text{traza}(At)}. \tag{9}$$

Demostración rápida de que e^{At} es invertible: *Simplemente reconozca e^{-At} como su inversa.*

Esta invertibilidad es fundamental para las ecuaciones diferenciales. Si n soluciones son linealmente independientes en $t = 0$, *son linealmente independientes por siempre*. Si los vectores iniciales son $v_1, \ldots, v_n$, entonces las soluciones $e^{At}v$ pueden escribirse en una matriz:

$$\begin{bmatrix} e^{At}v_1 & \cdots & e^{At}v_n \end{bmatrix} = e^{At}\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}.$$

El determinante del miembro izquierdo es el *wronskiano*. Éste nunca es cero, ya que es el producto de dos determinantes diferentes de cero. Las dos matrices en el miembro derecho son invertibles.

Observación No todas las ecuaciones diferenciales se presentan como un sistema de primer orden $du/dt = Au$. Es posible empezar, a partir de una simple ecuación de orden superior, como $y''' - 3y'' + 2y' = 0$. Para convertir lo anterior en un sistema de 3 por 3, se introducen $v = y'$ y $w = v'$ como incógnitas adicionales, junto con y misma. Luego, estas dos ecuaciones se combinan con la original para obtener $u' = Au$:

$$\begin{aligned} y' &= v \\ v' &= w \\ w' &= 3w - 2v \end{aligned} \qquad \text{o bien,} \quad u' = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & 3 \end{bmatrix}\begin{bmatrix} y \\ v \\ w \end{bmatrix} = Au.$$

De nuevo, se ha regresado a un sistema de primer orden. El problema puede resolverse de dos formas. En un curso de ecuaciones diferenciales, $y = e^{\lambda t}$ se sustituiría por $y''' - 3y'' + 2y' = 0$:

$$(\lambda^3 - 3\lambda^2 + 2\lambda)e^{\lambda t} = 0 \quad \text{o bien,} \quad \lambda(\lambda - 1)(\lambda - 2)e^{\lambda t} = 0. \tag{10}$$

Las tres soluciones exponenciales puras son $y = e^{0t}$, $y = e^t$, y $y = e^{2t}$. No está implicado ningún vector característico. En un curso de álgebra lineal, se encuentran los valores característicos de A:

$$\det(A - \lambda I) = \begin{bmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 0 & -2 & 3-\lambda \end{bmatrix} = -\lambda^3 + 3\lambda^2 - 2\lambda = 0. \tag{11}$$

¡Las ecuaciones (10) y (11) son las mismas! Aparecen los mismos tres exponentes: $\lambda = 0$, $\lambda = 1$, y $\lambda = 2$. Esta es una regla general que hace consistentes a los dos métodos; las razones de crecimiento de las soluciones permanecieron fijas cuando las ecuaciones cambiaron de forma. Nos parece que resolver la ecuación de tercer orden es más rápido.

La importancia física de $du/dt = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$ es fácil de explicar, y al mismo tiempo es realmente importante. Esta ecuación diferencial describe un proceso de *difusión*.

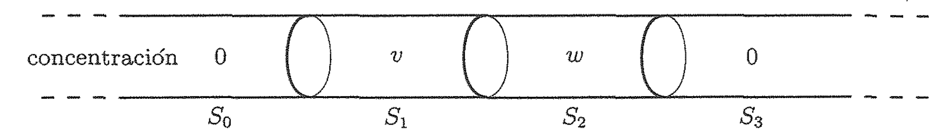


Figura 5.1 Modelo de difusión entre cuatro segmentos.

Divida un tubo infinito en cuatro segmentos (véase la figura 5.1). En el instante $t = 0$, los segmentos de en medio contienen concentraciones $v(0)$ y $w(0)$ de un producto químico. ***En cada instante t, la razón de difusión entre dos segmentos adyacentes es la diferencia en concentraciones.*** Dentro de cada segmento, la concentración permanece uniforme (cero en los segmentos infinitos). El proceso es continuo en el tiempo pero discreto en el espacio; las incógnitas son $v(t)$ y $w(t)$ en los dos segmentos interiores S_1 y S_2.

La concentración $v(t)$ en S_1 cambia de dos formas. Hay una difusión hacia S_0, y hacia o fuera de S_2. La razón de cambio neta es dv/dt, y dw/dt es semejante:

$$\textbf{Caudal hacia } S_1 \quad \frac{dv}{dt} = (w - v) + (0 - v)$$

$$\textbf{Caudal hacia } S_2 \quad \frac{dw}{dt} = (0 - w) + (v - w).$$

Esta ley de difusión coincide exactamente con nuestro ejemplo $du/dt = Au$:

$$u = \begin{bmatrix} v \\ w \end{bmatrix} \quad \text{y} \quad \frac{du}{dt} = \begin{bmatrix} -2v + w \\ v - 2w \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}u.$$

Los valores característicos -1 y -3 rigen la solución, y proporcionan la razón de que se desintegra la concentración, y λ_1 es el más importante porque sólo un conjunto excepcional de condiciones puede conducir a "superdesintegración" a la razón e^{-3t}. De hecho, estas condiciones deben provenir del vector característico $(1, -1)$. Si el experimento sólo admite concentraciones no negativas, es imposible y la razón limitante debe ser e^{-t}. La solución que se desintegra a esta razón más lenta, corresponde al vector característico $(1, 1)$. En consecuencia, las dos concentraciones se vuelven casi iguales (lo cual es típico de la difusión) cuando $t \to \infty$.

Un comentario adicional sobre este ejemplo: Se trata de una aproximación discreta, con sólo dos incógnitas, para la difusión continua descrita por esta ecuación diferencial parcial:

$$\textbf{Ecuación del calor} \qquad \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}.$$

Esa ecuación del calor es aproximada, al dividir el tubo en segmentos cada vez más pequeños, de longitud $1/N$. El sistema discreto con N incógnitas está regido por

$$\frac{d}{dt}\begin{bmatrix} u_1 \\ \cdot \\ \cdot \\ u_N \end{bmatrix} = \begin{bmatrix} -2 & 1 & & \\ 1 & -2 & \cdot & \\ & \cdot & \cdot & 1 \\ & & 1 & -2 \end{bmatrix}\begin{bmatrix} u_1 \\ \cdot \\ \cdot \\ u_N \end{bmatrix} = Au. \tag{12}$$

Esta es la matriz en diferencias finitas con el patrón 1, −2, 1. El miembro derecho de Au tiende a la segunda derivada d^2u/dx^2, luego que un factor de escala N^2 se presenta del problema de flujo. En el límite cuando $N \to \infty$, se llega a la ***ecuación del calor*** $\partial u/\partial t = \partial^2 u/\partial x^2$. Sus soluciones siguen siendo combinaciones de exponenciales puras, aunque ahora hay una infinidad. En vez de vectores característicos de $Ax = \lambda x$, se tienen *funciones características* de $d^2u/dx^2 = \lambda u$. Éstas son $u(x) = \operatorname{sen} n\pi x$ con $\lambda = -n^2\pi^2$. Así, la solución de la ecuación del calor es

$$u(t) = \sum_{n=1}^{\infty} c_n e^{-n^2\pi^2 t} \operatorname{sen} n\pi x.$$

Las constantes c_n están determinadas por la condición inicial. La novedad es que los vectores característicos son funciones de $u(x)$, ya que el problema es continuo y no discreto.

Estabilidad de las ecuaciones diferenciales

Así como para las ecuaciones en diferencias, los valores característicos deciden cómo se comporta $u(t)$ cuando $t \to \infty$. En la medida en que sea posible diagonalizar a A, hay n soluciones exponenciales puras de la ecuación diferencial, y cualquier solución específica $u(t)$ es alguna combinación

$$u(t) = Se^{\Lambda t}S^{-1}u_0 = c_1 e^{\lambda_1 t}x_1 + \cdots + c_n e^{\lambda_n t}x_n.$$

La estabilidad está regida por estos factores $e^{\lambda_i t}$. Si todos ellos tienden a cero, entonces $u(t)$ tiende a cero; si todos están acotados, entonces $u(t)$ permanece acotado; si uno de ellos se infla, entonces excepto por condiciones iniciales muy especiales la solución se infla. Además, el tamaño de $e^{\lambda t}$ sólo depende de la parte real de λ. ***Lo único que rige la estabilidad son las partes reales de los valores característicos:*** si $\lambda = a + ib$, entonces

$$e^{\lambda t} = e^{at}e^{ibt} = e^{at}(\cos bt + i \operatorname{sen} bt) \quad \text{y la magnitud es } |e^{\lambda t}| = e^{at}.$$

Lo anterior se desintegra para $a < 0$, es constante para $a = 0$, y explota para $a > 0$. La parte imaginaria produce oscilaciones, aunque la amplitud proviene de la parte real.

5M La ecuación diferencial $du/dt = Au$ es
estable y $e^{At} \to 0$ siempre que toda $\operatorname{Re} \lambda_i < 0$,
neutralmente estable cuando toda $\operatorname{Re} \lambda_i \leq 0$ y $\operatorname{Re} \lambda_1 = 0$, e
inestable y e^{At} no está acotado si cualquier valor característico tiene $\operatorname{Re} \lambda_i > 0$.

En algunos textos, la condición $\operatorname{Re} \lambda < 0$ se denomina estabilidad *asintótica*, porque garantiza desintegración para grandes instantes t. Nuestro razonamiento dependía de contar con n soluciones exponenciales puras, pero incluso si A no es diagonalizable (y haya términos como $te^{\lambda t}$) el resultado sigue siendo cierto: ***todas las soluciones tienden a cero si y sólo si todos los valores característicos tienen*** $\operatorname{Re} \lambda < 0$.

La estabilidad es especialmente fácil de decidir para un sistema de 2 por 2 (que es muy común en aplicaciones). La ecuación es

$$\frac{du}{dt} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} u,$$

y es necesario saber cuándo, ambos valores característicos de esta matriz tienen partes reales negativas. (Observe nuevamente que los valores característicos pueden ser números complejos). Las pruebas de estabilidad son

$\operatorname{Re} \lambda_1 < 0$	***la traza a + d debe ser negativa.***
$\operatorname{Re} \lambda_2 < 0$	***el determinante ad − bc debe ser positivo.***

Cuando los valores característicos son reales, estas pruebas garantizan que son negativos. Su producto es el determinante; es positivo cuando los valores característicos tienen el mismo signo. Su suma es la traza; es negativa cuando ambos valores característicos son negativos.

Cuando los valores característicos son un par complejo $x \pm iy$, las pruebas siguen teniendo éxito. La traza es su suma $2x$ (que es < 0) y el determinante es $(x + iy)(x - iy) = x^2 + y^2 > 0$. En la figura 5.2 se muestra el cuadrante estable, la traza < 0 y el determinante > 0. También se muestra la línea fronteriza parabólica entre los valores característicos reales y complejos. La razón de la parábola se encuentra en la ecuación cuadrática para los valores característicos:

$$\det \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix} = \lambda^2 - (\text{traza})\lambda + (\det) = 0. \tag{13}$$

La fórmula cuadrática para λ conduce a la parábola $(\text{traza})^2 = 4(\det)$:

$$\lambda_1 \ \text{ y } \ \lambda_2 = \frac{1}{2}\left[\text{traza} \pm \sqrt{(\text{traza})^2 - 4(\det)}\,\right]. \tag{14}$$

Arriba de la parábola, el número en el radicando es negativo, de modo que λ no es real. Sobre la parábola, la raíz cuadrada es cero y λ es repetido. Debajo de la parábola, las raíces cuadradas son reales. *Toda matriz simétrica tiene valores característicos reales*, ya que si $b = c$, entonces

$$(\text{traza})^2 - 4(\det) = (a + d)^2 - 4(ad - b^2) = (a - d)^2 + 4b^2 \geq 0.$$

Para valores característicos complejos, b y c tienen signos opuestos y son suficientemente grandes.

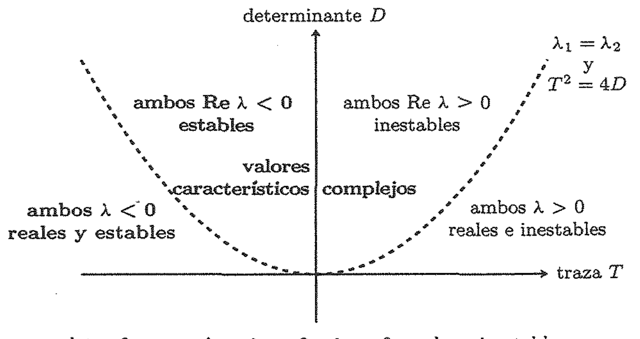


Figura 5.2 Regiones de estabilidad e inestabilidad para una matriz de 2 por 2.

Ejemplo 2 Uno de cada cuadrante: sólo #2 es estable:

$$\begin{bmatrix}1&0\\0&2\end{bmatrix}\quad\begin{bmatrix}-1&0\\0&-2\end{bmatrix}\quad\begin{bmatrix}1&0\\0&-2\end{bmatrix}\quad\begin{bmatrix}-1&0\\0&2\end{bmatrix}$$

Sobre las fronteras del segundo cuadrante, la ecuación es neutralmente estable. Sobre el eje horizontal, un valor característico es cero (porque el determinante es $\lambda_1\lambda_2 = 0$). Sobre el eje vertical arriba del eje, ambos valores característicos son puramente imaginarios (porque la traza es cero). Cruzando estos ejes están las dos formas en que se pierde la estabilidad.

El caso de n por n es más difícil. Una prueba para Re $\lambda_i < 0$ se debe a Routh y Hurwitz, quienes encontraron una serie de desigualdades sobre los elementos a_{ij}. No creo que este método sea demasiado bueno para una matriz grande; quizá la computadora encuentre los valores característicos con mayor certidumbre de los que puede probar tales desigualdades. La idea de Liapunov fue encontrar una *matriz de ponderación* W *tal que la longitud ponderada* $\|Wu(t)\|$ *siempre es decreciente*. Si existe una W así, entonces $\|Wu\|$ decrece de manera estable a cero, y después de algunos altibajos u también debe llegar ahí (estabilidad). El verdadero valor del método de Liapunov, se encuentra para una ecuación no lineal; entonces la estabilidad puede demostrarse sin necesidad de conocer una fórmula para $u(t)$.

Ejemplo 3 $du/dt = \begin{bmatrix}0&-1\\1&0\end{bmatrix}u$ manda a $u(t)$ alrededor de un círculo, empezando desde $u(0) = (1, 0)$.

Debido a que traza $= 0$ y det $= 1$, se tienen valores característicos puramente imaginarios:

$$\begin{bmatrix}-\lambda&-1\\1&-\lambda\end{bmatrix} = \lambda^2+1=0 \quad\text{de modo que}\quad \lambda = +i \quad\text{y}\quad -i.$$

Los vectores característicos son $(1, -i)$ y $(1, i)$, y la solución es

$$u(t) = \frac{1}{2}e^{it}\begin{bmatrix}1\\-i\end{bmatrix} + \frac{1}{2}e^{-it}\begin{bmatrix}1\\i\end{bmatrix}.$$

Lo anterior es correcto pero no es elegante. Al sustituir $\cos t \pm i \operatorname{sen} t$ para e^{it} y e^{-it}, *vuelven a aparecer números reales*: la solución circulante es $u(t) = (\cos t, \operatorname{sen} t)$.

Empezando desde una $u(0) = (a, b)$ distinta, la solución $u(t)$ termina como

$$u(t) = \begin{bmatrix}a\cos t - b\operatorname{sen} t\\ b\cos t + a\operatorname{sen} t\end{bmatrix} = \begin{bmatrix}\cos t & -\operatorname{sen} t\\ \operatorname{sen} t & \cos t\end{bmatrix}\begin{bmatrix}a\\b\end{bmatrix}. \qquad (15)$$

¡Aquí se tiene algo importante! La última matriz multiplica a $u(0)$, por lo que debe ser la exponencial e^{At}. (Recuerde que $u(t) = e^{At}u(0)$.) Esta matriz de senos y cosenos es nuestro ejemplo principal de una *matriz ortogonal*. La longitud de las columnas es 1, su producto interno es cero, y se tiene una confirmación de un hecho maravilloso:

Si A es simétrica sesgada ($A^T = -A$) entonces e^{At} es una matriz ortogonal.

$A^T = -A$ proporciona un sistema conservador. En amortiguamiento o difusión no se pierde energía:

$$A^T = -A, \qquad (e^{At})^T = e^{-At}, \qquad \text{y} \qquad \|e^{At}u(0)\| = \|u(0)\|.$$

La última ecuación expresa una propiedad esencial de las matrices ortogonales. Cuando multiplican a un vector, la longitud no cambia. El vector $u(0)$ simplemente rota, y eso describe la solución de $du/dt = Au$: *se desplaza alrededor de un círculo*.

En este caso bastante poco común, e^{At} también puede reconocerse directamente, a partir de la serie infinita.

Observe que $A = \begin{bmatrix}0&-1\\1&0\end{bmatrix}$ tiene $A^2 = -I$, lo cual se aplica en la serie para e^{At}:

$$I + At + \frac{(At)^2}{2} + \frac{(At)^3}{6} + \cdots = \begin{bmatrix}\left(1-\frac{t^2}{2}+\cdots\right) & \left(-t+\frac{t^3}{6}-\cdots\right)\\ \left(t-\frac{t^3}{6}+\cdots\right) & \left(1-\frac{t^2}{2}+\cdots\right)\end{bmatrix}$$

$$= \begin{bmatrix}\cos t & -\operatorname{sen} t\\ \operatorname{sen} t & \cos t\end{bmatrix}$$

Ejemplo 4 La ecuación de difusión es estable: $A = \begin{bmatrix}-2&1\\1&-2\end{bmatrix}$ tiene $\lambda = -1$ y $\lambda = -3$.

Ejemplo 5 Si los segmentos infinitos se cierran, nada puede escapar:

$$\frac{du}{dt} = \begin{bmatrix}-1&1\\1&-1\end{bmatrix}u \quad\text{o bien.}\quad \begin{aligned} dv/dt &= w - v\\ dw/dt &= v - w.\end{aligned}$$

Este es un *proceso de Markov continuo*. En vez de moverse cada año, las partículas se mueven a cada instante. Su número total $v + w$ es constante. Esto se debe al hecho de sumar dos ecuaciones en el miembro derecho: la derivada de $v + w$ es cero.

Una matriz de Markov discreta tiene la suma de sus columnas igual a $\lambda_{\text{máx}} = 1$. Una matriz de Markov *continua*, para ecuaciones diferenciales, tiene la suma de sus columnas igual a $\lambda_{\text{máx}} = 0$. A es una matriz de Markov discreta si y sólo si $B = A - I$ es una matriz de Markov continua. El estado estacionario para ambas es el vector característico para $\lambda_{\text{máx}}$. Se multiplica por $1^k = 1$ en ecuaciones en diferencias y por $e^{0t} = 1$ en ecuaciones diferenciales, y no se mueve.

En el ejemplo, el estado estacionario tiene $v = w$.

Ejemplo 6 En ingeniería nuclear, un reactor se denomina *crítico* cuando es neutralmente estable; la fisión equilibra la desintegración. Una fisión más lenta lo hace más estable, o *subcrítico*, y finalmente se desmantela. La fisión inestable es una bomba.

Ecuaciones de segundo orden

Las leyes de difusión llevaron a un sistema de primer orden $du/dt = Au$. Así lo hacen también muchas otras aplicaciones en química, biología, y otras áreas, aunque la ley más importante de la física no lo hace. Se trata de la *ley de Newton* $F = ma$, y la aceleración a es una segunda derivada. Los términos inerciales producen ecuaciones de segundo orden (es necesario resolver $d^2u/dt^2 = Au$ en vez de $du/dt = Au$), y el objetivo es entender cómo este giro a segundas derivadas modifica la solución.* En álgebra lineal es opcional, pero no es así en física.

*Las cuartas derivadas también son posibles, en la flexión de vigas, aunque la naturaleza parece resistir si se pasa a derivadas superiores a aquéllas.

La comparación es perfecta si se preserva la misma A:

$$\frac{d^2u}{dt^2} = Au = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} u. \tag{16}$$

Dos condiciones iniciales inician el sistema: el "desplazamiento" $u(0)$ y la "velocidad" $u'(0)$. Para cumplir estas condiciones, hay $2n$ soluciones exponenciales puras.

Suponga que se usa ω en vez de λ, y que estas soluciones especiales se escriben como $u = e^{i\omega t}x$. Al sustituir esta exponencial en la ecuación diferencial, debe cumplir

$$\frac{d^2}{dt^2}(e^{i\omega t}x) = A(e^{i\omega t}x), \quad \text{o bien,} \quad -\omega^2 x = Ax. \tag{17}$$

El vector x debe ser un vector característico de A, exactamente como antes. Ahora el valor característico correspondiente es $-\omega^2$, de modo que la frecuencia ω está relacionada con la razón de desintegración λ mediante la ley $-\omega^2 = \lambda$. Toda solución especial $e^{\lambda t}x$ de la ecuación de primer orden lleva a *dos* soluciones especiales $e^{i\omega t}x$ de la ecuación de segundo orden, y los dos exponentes son $\omega = \pm\sqrt{-\lambda}$. Lo anterior sólo fracasa cuando $\lambda = 0$, que sólo tiene una raíz cuadrada: si el vector característico es x, entonces las dos soluciones especiales son x y tx.

Para una matriz de difusión genuina, todos los valores característicos λ son negativos y todas las frecuencias ω son reales: *la difusión pura se transforma en oscilación pura*. Los factores $e^{i\omega t}$ producen estabilidad neutra, la solución no crece ni se desintegra, y la energía total permanece precisamente constante. Simplemente continúa pasando alrededor del sistema. La solución general de $d^2u/dt^2 = Au$, si A tiene valores característicos negativos $\lambda_1, \ldots, \lambda_n$ y si $\omega_j = \sqrt{-\lambda_j}$, es

$$u(t) = (c_1 e^{i\omega_1 t} + d_1 e^{-\omega_1 t})x_1 + \cdots + (c_n e^{i\omega_n t} + d_n e^{-i\omega_n t})x_n. \tag{18}$$

Como siempre, las constantes se encuentran a partir de las condiciones iniciales. Esto es más fácil de hacer (a expensas de una fórmula adicional) cambiando de exponenciales oscilatorias a las funciones seno y coseno más conocidas:

$$u(t) = (a_1 \cos\omega_1 t + b_1 \operatorname{sen}\omega_1 t)x_1 + \cdots + (a_n \cos\omega_n t + b_n \operatorname{sen}\omega_n t)x_n. \tag{19}$$

El desplazamiento inicial $u(0)$ es fácil de mantener separado: $t = 0$ significa que sen $\omega t = 0$ y cos $wt = 1$, dejando sólo

$$u(0) = a_1x_1 + \cdots + a_nx_n, \quad \text{o bien,} \quad u(0) = Sa, \quad \text{o bien,} \quad a = S^{-1}u(0).$$

Luego de diferenciar $u(t)$ y hacer $t = 0$, las bs se determinan por la velocidad inicial: $u'(0) = b_1\omega_1x_1 + \cdots + b_n\omega_nx_n$. La ecuación se resuelve al sustituir las as y las bs en la fórmula para $u(t)$.

La matriz $A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$ tiene $\lambda_1 = -1$ y $\lambda_2 = -3$. Las frecuencias son $\omega_1 = 1$ y $\omega_2 = \sqrt{3}$. Si el sistema inicia desde el reposo, $u'(0) = 0$, los términos en b sen ωt desaparecen:

$$\textbf{Solución de } u(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad u(t) = \frac{1}{2}\cos t \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2}\cos\sqrt{3}t \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Físicamente, dos masas están unidas entre sí y a paredes estacionarias mediante tres resortes idénticos (véase la figura 5.3). La primera masa se mantiene a $v(0) = 1$, la segunda se mantiene a $w(0) = 0$ y en $t = 0$ se libera. Su movimiento $u(t)$ se convierte en un promedio de dos oscilaciones puras, correspondientes a los vectores característicos. En el primer modo $x_1 = (1, 1)$, las masas se mueven juntas y el resorte en el medio jamás se es-

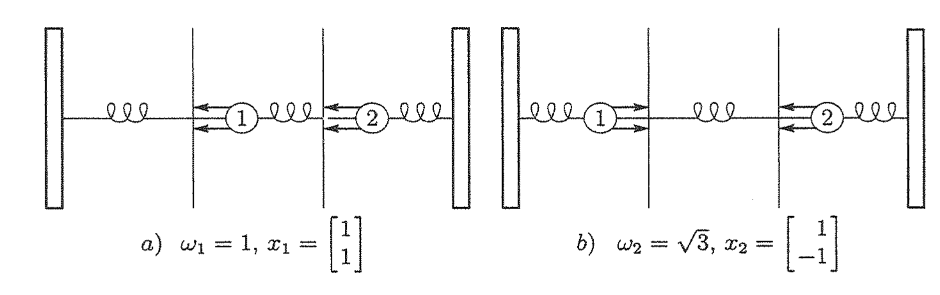


Figura 5.3 Los modos de oscilación rápida y lenta.

tira (véase la figura 5.3a). La frecuencia $\omega_1 = 1$ es la misma que para un solo resorte y una sola masa. En el modo más rápido $x_2 = (1, -1)$ con frecuencia $\sqrt{3}$, las masas se mueven de manera opuesta pero a la misma velocidad. La solución general es una combinación de estos dos modos normales. La solución particular es la mitad de cada una.

A medida que transcurre el tiempo, el movimiento es "casi periódico". Si la razón ω_1/ω_2 fuese una fracción como 2/3, las masas terminarían por volver a $u(0) = (1, 0)$ y empezarían de nuevo. Una combinación de sen $2t$ y sen $3t$ tendría un periodo de 2π. Sin embargo, $\sqrt{3}$ es irracional. Lo mejor que puede decirse es que las masas estarán *arbitrariamente próximas* a (1, 0) y también a (0, 1). Como una bola de billar que rebota por siempre en una mesa perfectamente lisa, la energía total es fija. Tarde o temprano las masas llegan cerca de cualquier estado con esta energía.

De nuevo es imposible dejar el problema sin establecer un paralelismo con el caso continuo. A medida en que las masas y los resortes discretos se unen en una varilla sólida, las "segundas diferencias" dadas por la matriz A 1, -2, 1 se convierten en segundas derivadas. Este límite es descrito por la célebre ***ecuación de onda*** $\partial^2u/\partial t^2 = \partial^2u/\partial x^2$.

Conjunto de problemas 5.4

1. Siga el primer ejemplo de esta sección, para encontrar los valores característicos y los vectores característicos, así como la exponencial e^{At}, para

$$A = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}.$$

2. Para la matriz previa, escriba la solución general de $du/dt = Au$, así como la solución específica que coincida con $u(0) = (3, 1)$. ¿Cuál es el *estado estacionario* cuando $t \to \infty$? (Este es un proceso de Markov continuo; $\lambda = 0$ en una ecuación diferencial corresponde a $\lambda = 1$ en una ecuación en diferencias, ya que $e^{0t} = 1$.)

3. Suponga que la dirección temporal se invierte para obtener la matriz $-A$:

$$\frac{du}{dt} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} u \quad \text{con} \quad u_0 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

Encuentre $u(t)$ y demuestre que *se infla* en vez de desintegrarse cuando $t \to \infty$. (La difusión es irreversible, y la ecuación del calor no puede aplicarse en retrospectiva.)

4. Si P es una matriz proyección, a partir de la serie infinita demuestre que

$$e^P \approx I + 1.718P.$$

5. Una matriz diagonal como $\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ satisface la regla de costumbre $e^{\Lambda(t+T)} = e^{\Lambda t}e^{\Lambda T}$ ya que la regla se cumple para cada elemento diagonal.

a) Explique por qué $e^{A\ t+T)} = e^{At}e^{AT}$, usando la fórmula $e^{At} = Se^{\Lambda t}S^{-1}$.

b) Demuestre que $e^{A+B} = e^A e^B$ *no se cumple* para matrices a partir del ejemplo

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \quad \text{(utilice la serie para } e^A \text{ y } e^B\text{)}$$

6. La ecuación de orden superior $y'' + y = 0$ puede escribirse como un sistema de primer orden si la velocidad y' se introduce como otra incógnita:

$$\frac{d}{dt}\begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} y' \\ y'' \end{bmatrix} = \begin{bmatrix} y' \\ -y \end{bmatrix}.$$

Si ésta es $du/dt = Au$, ¿cuál es la matriz A de 2 por 2? Encuentre sus valores característicos y sus vectores característicos, y calcule la solución que empieza desde $y(0) = 2$, $y'(0) = 0$.

7. Transforme $y'' = 0$ a un sistema $du/dt = Au$ de primer orden:

$$\frac{d}{dt}\begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} y' \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} y \\ y' \end{bmatrix}.$$

Esta matriz A de 2 por 2 sólo tiene un vector característico y no es posible diagonalizarla. Calcule e^{At}, a partir de la serie $I + At + \cdots$ y escriba la solución $e^{At}u(0)$ empezando desde $y(0) = 3$, $y'(0) = 4$. Compruebe que su (y, y') satisface $y'' = 0$.

8. Suponga que la población de conejos r, y la población de lobos w están regidas por

$$\frac{dr}{dt} = 4r - 2w$$

$$\frac{dw}{dt} = r + w.$$

a) Este sistema es ¿estable, neutralmente estable o inestable?

b) Si inicialmente $r = 300$ y $w = 200$, ¿cuáles son las poblaciones en el instante t?

c) Al cabo de bastante tiempo, ¿cuál es la proporción de conejos a lobos?

9. Decida la estabilidad de $u' = Au$ para las siguientes matrices:

a) $A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$. *b*) $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$.

c) $A = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}$. *d*) $A = \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}$.

10. Decida la estabilidad o inestabilidad de $dv/dt = w$, $dw/dt = v$. ¿Existe una solución que se desintegre?

11. Con base en su traza y su determinante, ¿en qué instante t las siguientes matrices cambian entre estables con valores característicos reales, estables con valores característicos complejos e inestables?

$$A_1 = \begin{bmatrix} 1 & -1 \\ t & -1 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 0 & 4-t \\ 1 & -2 \end{bmatrix}, \qquad A_3 = \begin{bmatrix} t & -1 \\ 1 & t \end{bmatrix}.$$

12. Encuentre los valores característicos y los vectores característicos para

$$\frac{du}{dt} = Au = \begin{bmatrix} 0 & 3 & 0 \\ -3 & 0 & 4 \\ 0 & -4 & 0 \end{bmatrix} u.$$

¿Por qué se sabe, sin necesidad de calcular, que e^{At} es una matriz ortogonal y que $\|u(t)\|^2 = u_1^2 + u_2^2 + u_3^2$ es constante?

13. Para la ecuación simétrica sesgada

$$\frac{du}{dt} = Au = \begin{bmatrix} 0 & c & -b \\ -c & 0 & a \\ b & -a & 0 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix},$$

a) Escriba u_1', u_2', u_3' y confirme que $u_1'u_1 + u_2'u_2 + u_3'u_3 = 0$.

b) Deduzca que la longitud $u_1^2 + u_2^2 + u_3^2$ es una constante.

c) Encuentre los valores característicos de A.

La solución gira alrededor del eje $w = (a, b, c)$, porque Au es el "producto cruz" $u \times w$, que es perpendicular a u y w.

14. ¿Cuáles son los valores característicos λ y las frecuencias ω, así como la solución general de la siguiente ecuación?

$$\frac{d^2u}{dt^2} = \begin{bmatrix} -5 & 4 \\ 4 & -5 \end{bmatrix} u.$$

15. Resuelva la ecuación de segundo orden

$$\frac{d^2u}{dt^2} = \begin{bmatrix} -5 & -1 \\ -1 & -5 \end{bmatrix} u \quad \text{con} \quad u(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{y} \quad u'(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

16. En la mayor parte de las aplicaciones, la ecuación de segundo orden se ve como $Mu'' + Ku = 0$, con una *matriz masa* que multiplica a las segundas derivadas. Sustituya la exponencial pura $u = e^{i\omega t}x$ y encuentre el "problema de valores característicos generalizado" que es necesario resolver para la frecuencia ω y el vector x.

17. Con una matriz fricción F en la ecuación $u'' + Fu' - Au = 0$, sustituya una exponencial pura $u = e^{\lambda t}x$, y encuentre un problema de valor característico cuadrático para λ.

18. Para la ecuación (16) en el texto, con $\omega = 1\ \sqrt{3}$, encuentre el movimiento si la primera masa se golpea en $t = 0$; $u(0) = (0, 0)$ y $u'(0) = (1, 0)$.

19. Toda matriz de 2 por 2 con traza cero puede escribirse como

$$A = \begin{bmatrix} a & b+c \\ b-c & -a \end{bmatrix}.$$

Demuestre que sus valores característicos son reales exactamente cuando $a^2 + b^2 \geq c^2$.

20. Por sustitución hacia atrás o calculando los vectores característicos, resuelva

$$\frac{du}{dt} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 3 & 6 \\ 0 & 0 & 4 \end{bmatrix} u \quad \text{con} \quad u(0) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

21. Encuentre los λs y los xs de modo que $e^{\lambda t}x$ resuelva

$$\frac{du}{dt} = \begin{bmatrix} 4 & 3 \\ 0 & 1 \end{bmatrix} u.$$

¿Qué combinación $u = c_1 e^{\lambda_1 t} x_1 + c_2 e^{\lambda_2 t} x_2$ empieza desde $u(0) = (5, -2)$?

22. Resuelva el problema 21 para $u(t) = (y(t), z(t))$ por sustitución hacia atrás:

Primero se resuelve $\frac{dz}{dt} = z$, empezando con $z(0) = -2$.

Luego se resuelve $\frac{dy}{dt} = 4y + 3z$, empezando con $y(0) = 5$.

La solución para y es una combinación de e^{4t} y e^t.

23. Encuentre A para cambiar $y'' = 5y' + 4y$ en una ecuación vectorial para $u(t) = (y(t), y'(t))$:

$$\frac{du}{dt} = \begin{bmatrix} y' \\ y'' \end{bmatrix} = \begin{bmatrix} & \\ & \end{bmatrix} \begin{bmatrix} y \\ y' \end{bmatrix} = Au.$$

¿Cuáles son los valores característicos de A? Encuéntrelos también por sustitución de $y = e^{\lambda t}$ en la ecuación escalar $y'' = 5y' + 4y$.

24. Entre dos habitaciones con aforo para $v(0) = 30$ personas y $w(0) = 10$ personas se abre una puerta. El movimiento entre las habitaciones es proporcional a la diferencia $v - w$:

$$\frac{dv}{dt} = w - v \qquad \text{y} \qquad \frac{dw}{dt} = v - w.$$

Demuestre que el total $v + w$ es constante (40 personas). Encuentre la matriz en $du/dt = Au$, así como sus valores característicos y vectores característicos. ¿Cuáles son v y w en $t = 1$?

25. Invierta la difusión de personas en el problema 24 a $du/dt = -Au$:

$$\frac{dv}{dt} = v - w \qquad \text{y} \qquad \frac{dw}{dt} = w - v.$$

El total $v + w$ sigue siendo constante. ¿Cómo cambian los λs ahora que A ha cambiado a $-A$? Sin embargo, demuestre que $v(t)$ tiende a infinito a partir de $v(0) = 30$.

26. La solución de $y'' = 0$ es una recta $y = C + Dt$. Transfórmela en una ecuación matricial:

$$\frac{d}{dt}\begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} y \\ y' \end{bmatrix} \text{ tiene la solución } \begin{bmatrix} y \\ y' \end{bmatrix} = e^{At}\begin{bmatrix} y(0) \\ y'(0) \end{bmatrix}.$$

Esta matriz A no puede diagonalizarse. Encuentre A^2 y calcule $e^{At} = I + At + \frac{1}{2}A^2t^2 + \cdots$. Multiplique su e^{At} por $(y(0), y'(0))$ para comprobar la recta $y(t) = y(0) + y'(0)t$.

27. Sustituya $y = e^{\lambda t}$ en $y'' = 6y' - 9y$ para demostrar que $\lambda = 3$ es una raíz repetida. Esto es un problema: se requiere una segunda solución después de e^{3t}. La ecuación matricial es

$$\frac{d}{dt}\begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -9 & 6 \end{bmatrix}\begin{bmatrix} y \\ y' \end{bmatrix}.$$

Demuestre que esta matriz tiene $\lambda = 3, 3$ y sólo una recta de vectores característicos. *Aquí también se tiene un problema.* Demuestre que la segunda solución es $y = te^{3t}$.

28. Imagine cómo escribir $my'' + by' + ky = 0$ como una ecuación vectorial $Mu' = Au$.

29. *a*) Encuentre dos funciones conocidas que resuelvan la ecuación $d^2y/dt^2 = -y$. ¿Cuál empieza con $y(0) = 1$ y $y'(0) = 0$?

b) La ecuación de segundo orden $y'' = -y$ produce una ecuación vectorial $u' = Au$:

$$u = \begin{bmatrix} y \\ y' \end{bmatrix} \qquad \frac{du}{dt} = \begin{bmatrix} y' \\ y'' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} y \\ y' \end{bmatrix} = Au.$$

Escriba $y(t)$ del inciso *a*) en $u(t) = (y, y')$. Esto resuelve nuevamente el problema 6.

30. Una solución particular de $du/dt = Au - b$ es $u_p = A^{-1}b$, si A es invertible. Las soluciones de $du/dt = Au$ dan u_n. Encuentre la solución completa $u_p + u_n$ de

$$a)\ \frac{du}{dt} = 2u - 8. \qquad b)\ \frac{du}{dt} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} u - \begin{bmatrix} 8 \\ 6 \end{bmatrix}.$$

31. Si c no es un valor característico de A, sustituya $u = e^{ct}v$, y encuentre v para resolver $du/dt = Au - e^{ct}b$. Esta $u = e^{ct}v$ es una solución particular. ¿Cómo fracasa cuando c es un valor característico?

32. Encuentre una matriz A para ilustrar cada una de las regiones inestables en la figura 5.2:

a) $\lambda_1 < 0$ y $\lambda_2 > 0$.
b) $\lambda_1 > 0$ y $\lambda_2 > 0$.
c) λs complejos con parte real $a > 0$.

Los problemas 33 a 41 son sobre la matriz exponencial e^{At}

33. Escriba cinco términos de la serie infinita para e^{At}. Tome la derivada t de cada término. Demuestre que tiene cuatro términos de Ae^{At}. Conclusión: $e^{At}u(0)$ resuelve $u' = Au$.

34. La matriz $B = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}$ tiene $B^2 = 0$. Encuentre e^{Bt}, a partir de una serie infinita (corta). Compruebe que la derivada de e^{Bt} es Be^{Bt}.

35. Empezando desde $u(0)$, la solución en el instante T es $e^{AT}u(0)$. Pase a un instante adicional t para llegar a $e^{At}(e^{AT}u(0))$. Esta solución en el instante $t + T$ también puede escribirse como ______. Conclusión: e^{At} multiplicado por e^{AT} es igual a ______.

36. Escriba $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ en la forma $S\Lambda S^{-1}$. Encuentre e^{At}, a partir de $Se^{\Lambda t}S^{-1}$.

37. Si $A^2 = A$, demuestre que la serie infinita produce $e^{At} = I + (e^t - 1)A$. Para $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ en el problema 36, con lo anterior se obtiene $e^{At} =$ ______.

38. En general, $e^A e^B$ es diferente de $e^B e^A$, y ambas son distintas de e^{A+B}. Compruebe esto, usando los problemas 36-37 y 34:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \qquad A + B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

39. Escriba $A = \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix}$ como $S\Lambda S^{-1}$. Multiplique $Se^{\Lambda t}S^{-1}$ para encontrar la matriz exponencial e^{At}. Compruebe que $e^{At} = I$ cuando $t = 0$.

40. Escriba $A = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}$ en la serie infinita para encontrar e^{At}. Primero calcule A^2:

$$e^{At} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} t & 3t \\ 0 & 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} & \\ & \end{bmatrix} + \cdots = \begin{bmatrix} e^t & \\ 0 & \end{bmatrix}.$$

41. Proporcione dos razones por las cuales la matriz exponencial e^{At} nunca es singular.
a) Escriba su inversa.
b) Escriba sus valores característicos. Si $Ax = \lambda x$, entonces $e^{At}x =$ ______x.

42. Encuentre una solución $x(t)$, $y(t)$ del primer sistema que se haga grande cuando $t \to \infty$. Para evitar esta inestabilidad, ¡un científico consideró intercambiar las dos ecuaciones!

$$\begin{aligned} dx/dt &= 0x - 4y \\ dy/dt &= -2x + 2y \end{aligned} \quad \text{se convierte en} \quad \begin{aligned} dy/dt &= -2x + 2y \\ dx/dt &= 0x - 4y. \end{aligned}$$

Así, la matriz $\begin{bmatrix} -2 & 2 \\ 0 & -4 \end{bmatrix}$ es estable. Tiene $\lambda < 0$. Comente esta locura.

43. A partir de esta solución general de $du/dt = Au$, encuentre la matriz A:

$$u(t) = c_1 e^{2t}\begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 e^{5t}\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

5.5 MATRICES COMPLEJAS

Ya no es posible trabajar sólo con vectores y matrices reales. En la primera parte de este libro, cuando el problema básico era $Ax = b$, la solución era real cuando A y b eran reales. Hubiera sido posible permitir números complejos, aunque eso no hubiera contribuido a nada. Ahora ya no es posible evitarlos. Una matriz real tiene coeficientes reales en det $(A - \lambda I)$, aunque los valores característicos (como en las rotaciones) pueden ser complejos.

A continuación se introduce el espacio $\mathbf{C}^n$ de vectores con n componentes *complejas*. La suma y la multiplicación matriciales obedecen las mismas reglas que antes. ***La longitud se calcula de otra forma.*** Como se hacía antes, el vector en $\mathbf{C}^2$ con componentes $(1, i)$ podía tener longitud cero: $1^2 + i^2 = 0$, lo cual no está bien. La longitud al cuadrado correcta es $1^2 + |i|^2 = 2$.

Este cambio a $\|x\|^2 + |x_1|^2 + \cdots + |x_n|^2$ obliga a toda una serie de cambios. El producto interno, la traspuesta, las definiciones de matrices simétrica y ortogonal, todo esto debe modificarse para números complejos. Las nuevas definiciones coinciden con las viejas cuando los vectores y matrices son reales. Estos cambios se presentan en una lista al final de la sección, y cada cuestión se explica en la medida en que se aborda.

Esa tabla es virtualmente un diccionario para pasar de real a complejo. Esperamos que sea de utilidad para el lector. En particular, se desea incursionar en las ***matrices simétricas y en las matrices hermitianas***: *¿Dónde están sus valores característicos, y qué hay de especial sobre sus vectores característicos*? Para efectos prácticos, estas son las cuestiones más importantes en la teoría de los valores característicos. De antemano se solicita atención para las respuestas:

1. ***Toda matriz simétrica (y toda matriz hermitiana) tiene valores característicos reales.***
2. ***Sus vectores característicos pueden escogerse de modo que sean ortonormales.***

Aunque parezca extraño, para demostrar que los valores característicos son reales se empieza con la posibilidad opuesta, que lleva a números complejos, vectores complejos y matrices complejas.

Números complejos y sus conjugados

Quizá el lector ya tenga experiencia con los números complejos; es fácil proporcionar un repaso. Los conceptos importantes son el *conjugado complejo* $\bar{x}$ y el *valor absoluto* $|x|$. Todo mundo sabe que, no importa qué sea i, si éste cumple la ecuación $i^2 = -1$. Se trata de un número imaginario puro, como lo son sus múltiplos ib; b es real. La suma $a + ib$ es un número complejo, que se grafica de forma natural en el plano complejo (véase la figura 5.4).

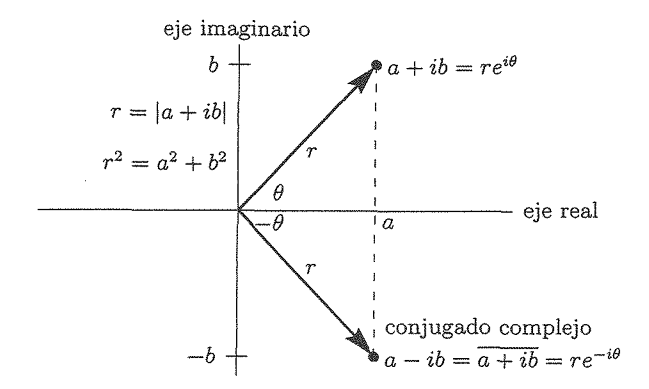


Figura 5.4 El plano complejo, con $a + ib = re^{i\theta}$ y su conjugado $a - ib = re^{-i\theta}$.

Los números reales a y los números imaginarios ib son los casos especiales de los números complejos; están sobre los ejes. La suma de dos números complejos es fácil:

$$\textbf{Suma de números complejos} \quad (a + ib) + (c + id) = (a + c) + i(b + d).$$

Al multiplicar $a + ib$ por $c + id$ se aplica la regla de que $i^2 = -1$:

$$\textbf{Multiplicación} \qquad \begin{aligned} (a + ib)(c + id) &= ac + ibc + iad + i^2 bd \\ &= (ac - bd) + i(bc + ad). \end{aligned}$$

El ***conjugado complejo*** de $a + ib$ es el número $a - ib$. *Se invierte el signo de las partes imaginarias.* Es la imagen especular a través del eje real; cualquier número real es su propio conjugado, ya que $b = 0$. El conjugado se denota por una barra o una estrella: $(a + ib)^* = \overline{a + ib} = a - ib$. Posee tres propiedades importantes:

1. El conjugado de un producto es igual al producto de los conjugados:

$$\overline{(a + ib)(c + id)} = (ac - bd) - i(bc + ad) = \overline{(a + ib)}\ \overline{(c + id)}. \qquad (1)$$

2. El conjugado de una suma es igual a la suma de los conjugados:

$$\overline{(a + c) + i(b + d)} = (a + c) - i(b + d) = \overline{(a + ib)} + \overline{(c + id)}.$$

3. Al multiplicar $a + ib$ por su conjugado $a - ib$ se obtiene un número real $a^2 + b^2$:

$$\textbf{Valor absoluto} \qquad (a + ib)(a - ib) = a^2 + b^2 = r^2. \qquad (2)$$

Esta distancia r es el ***valor absoluto*** $|a + ib| = \sqrt{a^2 + b^2}$.

Por último, la trigonometría relaciona los catetos a y b con la hipotenusa r mediante $a = r\cos\theta$ y $b = r\,\mathrm{sen}\,\theta$. Al combinar estas dos ecuaciones se llega a las coordenadas polares:

$$\textbf{Forma polar} \qquad a + ib = r(\cos\theta + i\,\mathrm{sen}\theta) = re^{i\theta}. \qquad (3)$$

El caso especial más importante es cuando $r = 1$. Así, $a + ib$ es $e^{i\theta} = \cos\theta + i\,\mathrm{sen}\theta$, y se encuentra sobre la ***circunferencia unitaria*** en el plano complejo. A medida que θ varía de 0 a 2π, este número $e^{i\theta}$ circula alrededor de cero a la distancia radial constante $|e^{i\theta}| = \sqrt{\cos^2\theta + \mathrm{sen}^2\theta} = 1$.

Ejemplo 1 $x = 3 + 4i$ multiplicado por su conjugado $\bar{x} = 3 - 4i$ es igual al valor absoluto al cuadrado:

$$x\bar{x} = (3 + 4i)(3 - 4i) = 25 = |x|^2 \quad \text{de modo que} \quad r = |x| = 5.$$

Para dividir entre $3 + 4i$, el numerador y el denominador se multiplican por su conjugado $3 - 4i$:

$$\frac{2+i}{3+4i} = \frac{2+i}{3+4i}\,\frac{3-4i}{3-4i} = \frac{10-5i}{25}.$$

En coordenadas polares, la multiplicación y la división son fáciles:

$re^{i\theta}$ multiplicado por $Re^{i\alpha}$ tiene valor absoluto rR y ángulo $\theta + \alpha$.
$re^{i\theta}$ dividido entre $Re^{i\alpha}$ tiene valor absoluto r/R y ángulo $\theta - \alpha$.

Longitudes y traspuestas en el caso complejo

Volvemos al álgebra lineal, y se efectúa la conversión de reales a complejos. Por definición, *el espacio vectorial complejo* $\mathbf{C}^n$ *contiene a todos los vectores x con n componentes complejas*:

$$\textbf{Vector complejo} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{con componentes} \quad x_j = a_j + ib_j.$$

La suma de los vectores x y y sigue haciéndose componente por componente. La multiplicación escalar cx ahora se realiza con números complejos c. Los vectores $v_1, \ldots, v_k$ son linealmente *dependientes* si alguna combinación no trivial proporciona $c_1v_1 + \cdots + c_kv_k = 0$; ahora los c_j deben ser complejos. Los vectores de coordenadas unitarios siguen estando en $\mathbf{C}^n$; siguen siendo independientes, y siguen siendo una base. En consecuencia, $\mathbf{C}^n$ es un espacio vectorial complejo de dimensión n.

En la nueva definición de longitud, cada x_j^2 se sustituye por su módulo $|x_j|^2$:

$$\textbf{Longitud al cuadrado} \qquad \|x\|^2 = |x_1|^2 + \cdots + |x_n|^2. \qquad (4)$$

Ejemplo 2 $x = \begin{bmatrix} 1 \\ i \end{bmatrix}$ y $\|x\|^2 = 2$; $y = \begin{bmatrix} 2+i \\ 2-4i \end{bmatrix}$ y $\|y\|^2 = 25$.

Para vectores reales, entre la longitud y el producto interno había una relación estrecha: $\|x\|^2 = x^{\mathrm{T}}x$. Se desea preservar esta relación. El producto interno debe modificarse para coincidir con la nueva definición de longitud, *de modo que en el producto interno se conjuga el primer vector*. Al sustituir x por $\bar{x}$, el ***producto interno se convierte en***

$$\textbf{Producto interno} \qquad \bar{x}^{\mathrm{T}}y = \bar{x}_1y_1 + \cdots + \bar{x}_ny_n. \qquad (5)$$

Si se toma el producto interno de $x = (1 + i, 3i)$ consigo mismo, se regresa a $\|x\|^2$:

$$\textbf{Longitud al cuadrado} \quad \bar{x}^{\mathrm{T}}x = \overline{(1+i)}(1+i) + \overline{(3i)}(3i) = 2 + 9 \quad \text{y} \quad \|x\|^2 = 11.$$

Observe que $\bar{y}^{\mathrm{T}}x$ es diferente de $\bar{x}^{\mathrm{T}}y$; es necesario vigilar el orden de los vectores.

Lo anterior deja un solo cambio más en la notación, resumiendo dos símbolos en uno. En vez de una barra para el conjugado y una T para la traspuesta, estos símbolos se combinan en la ***traspuesta conjugada***. Para vectores y matrices, un supraíndice H (o una estrella) combina ambas operaciones. Esta matriz $\bar{A}^{\mathrm{T}} = A^{\mathrm{H}} = A^*$ se denomina "*A* hermitiana":

$$\textbf{"A hermitiana"} \qquad A^{\mathrm{H}} = \bar{A}^{\mathrm{T}} \text{ tiene elementos } (A^{\mathrm{H}})_{ij} = \overline{A_{ji}}. \qquad (6)$$

El lector debe estar muy atento para distinguir esta denominación de la expresión "*A es* hermitiana", lo cual significa que *A es igual* a A^{H}. Si A es una matriz de m por n, entonces A^{H} es de n por m:

$$\textbf{Traspuesta conjugada} \qquad \begin{bmatrix} 2+i & 3i \\ 4-i & 5 \\ 0 & 0 \end{bmatrix}^{\mathrm{H}} = \begin{bmatrix} 2-i & 4+i & 0 \\ -3i & 5 & 0 \end{bmatrix}.$$

Este símbolo A^{H} otorga reconocimiento oficial al hecho de que, con elementos complejos, es muy raro que sólo se busque la traspuesta de A. Es la traspuesta *conjugada* A^{H} lo que se vuelve idóneo, y x^{H} es el vector renglón $[\bar{x}_1 \ \ldots \ \bar{x}_n]$.

5N 1. El producto interno de x y y es $x^{\mathrm{H}}y$. Los vectores ortogonales cumplen $x^{\mathrm{H}}y = 0$.
2. La longitud al cuadrado de x es $\|x\|^2 = x^{\mathrm{H}}x = |x_1|^2 + \cdots + |x_n|^2$.
3. Al conjugar $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$ se obtiene $(AB)^{\mathrm{H}} = B^{\mathrm{H}}A^{\mathrm{H}}$.

Matrices hermitianas

En capítulos previos se habló de matrices simétricas: $A = A^{\mathrm{T}}$. Con elementos complejos, es necesario ampliar el concepto de simetría. La generalización correcta no es hacia matrices que son iguales a su traspuesta, sino a ***matrices que son iguales a su traspuesta conjugada***. Estas son las matrices hermitianas, de las cuales un ejemplo típico es:

$$\textbf{Matriz hermitiana} \qquad A = \begin{bmatrix} 2 & 3-3i \\ 3+3i & 5 \end{bmatrix} = A^{\mathrm{H}}. \qquad (7)$$

Los elementos diagonales deben ser reales; la conjugación los deja sin cambio. Cada elemento fuera de la diagonal es compensado por su imagen especular a través de la diagonal principal, y $3 - 3i$ es el conjugado de $3 + 3i$. *En cada caso*, $a_{ij} = \overline{a_{ji}}$.

Nuestro objetivo principal es establecer tres propiedades básicas de las matrices hermitianas. Estas propiedades son igualmente válidas para las matrices simétricas. *Una matriz simétrica real ciertamente es hermitiana.* (Para matrices reales no hay diferencia entre A^{T} y A^{H}). ***Los valores característicos de A son reales***, como se demostrará a continuación.

Propiedad 1 Si $A = A^{\mathrm{H}}$, entonces para todos los vectores complejos x, el número $x^{\mathrm{H}}Ax$, es real.

Cada elemento de A contribuye a x^HAx. Intente el caso de 2 por 2 con $x = (u, v)$:

$$x^HAx = \begin{bmatrix}\bar{u} & \bar{v}\end{bmatrix}\begin{bmatrix}2 & 3-3i\\ 3+3i & 5\end{bmatrix}\begin{bmatrix}u\\ v\end{bmatrix}$$
$$= 2\bar{u}u + 5\bar{v}v + (3-3i)\bar{u}v + (3+3i)u\bar{v}$$
$$= \textbf{real} + \textbf{real} + (\textbf{suma de conjugados complejos}).$$

Para una demostración en general, $(x^HAx)^H$ es el conjugado de la matriz de 1 por 1 x^HAx, aunque en realidad se obtiene otra vez el mismo número: $(x^HAx)^H = x^HA^Hx^{HH} = x^HAx$. Por tanto, este número debe ser real.

Propiedad 2 Si $A = A^H$, todo valor característico es real.

Demostración Suponga que $Ax = \lambda x$. ***El truco consiste en multiplicar por*** x^H: $x^HAx = \lambda x^Hx$. Por la propiedad 1, el miembro izquierdo es real, y el miembro derecho $x^Hx = \|x\|^2$ es real y positivo, ya que $x \neq 0$. En consecuencia, $\lambda = x^HAx/x^Hx$ debe ser real. En el ejemplo se tiene $\lambda = 8$ y $\lambda = -1$:

$$|A - \lambda I| = \begin{vmatrix}2-\lambda & 3-3i\\ 3+3i & 5-\lambda\end{vmatrix} = \lambda^2 - 7\lambda + 10 - |3-3i|^2$$
$$= \lambda^2 - 7\lambda - 8 = (\lambda - 8)(\lambda + 1). \quad (8)$$

■

Nota Esta demostración de valores característicos reales es correcta para cualquier matriz real:

Demostración falsa $Ax = \lambda x$ proporciona $x^TAx = \lambda x^Tx$, de modo que $\lambda = \dfrac{x^TAx}{x^Tx}$ es real.

Debe haber una trampa: *El vector característico x podría ser complejo*. Es cuando $A = A^T$ que se tiene la certeza de que λ y x son reales. Más que eso, *los vectores característicos son perpendiculares*: $x^Ty = 0$ en el caso simétrico real y $x^Hy = 0$ en el caso hermitiano complejo.

Propiedad 3 Dos vectores característicos de una matriz simétrica real o de una matriz hermitiana, si provienen de valores característicos distintos, son ortogonales entre sí.

La demostración empieza con $Ax = \lambda_1 x$, $Ay = \lambda_2 y$, y $A = A^H$:

$$(\lambda_1 x)^H y = (Ax)^H y = x^HAy = x^H(\lambda_2 y). \quad (9)$$

Los números exteriores son $\lambda_1 x^Hy = \lambda_2 x^Hy$, ya que los λs son reales. Ahora se utiliza la hipótesis de que $\lambda_1 \neq \lambda_2$, *que obliga a concluir que* $x^Hy = 0$. En el ejemplo,

$$(A - 8I)x = \begin{bmatrix}-6 & 3-3i\\ 3+3i & -3\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}, \qquad x = \begin{bmatrix}1\\ 1+i\end{bmatrix}$$
$$(A + I)y = \begin{bmatrix}3 & 3-3i\\ 3+3i & 6\end{bmatrix}\begin{bmatrix}y_1\\ y_2\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}, \qquad y = \begin{bmatrix}1-i\\ -1\end{bmatrix}.$$

Estos dos vectores característicos son ortogonales:

$$x^Hy = \begin{bmatrix}1 & 1-i\end{bmatrix}\begin{bmatrix}1-i\\ -1\end{bmatrix} = 0.$$

Por supuesto, cualesquiera múltiplos x/α y y/β son igualmente válidos que los vectores característicos. MATLAB escoge $\alpha = \|x\|$ y $\beta = \|y\|$, de modo que x/α y y/β son vectores unitarios; los vectores característicos son normalizados para que su longitud sea 1. Ahora son *ortonormales*. Si estos vectores característicos se escogen como las columnas de S, entonces se tiene $S^{-1}AS = \Lambda$, como siempre. ***La matriz de diagonalización puede escogerse con columnas ortonormales cuando*** $A = A^H$.

En caso de que A sea real y simétrica, por la propiedad 2 sus valores característicos son reales. Por la propiedad 3, sus vectores característicos unitarios son ortogonales y también son reales; resuelven $(A - \lambda I)x = 0$. *Estos vectores característicos ortonormales van en una matriz ortogonal* Q,
con $Q^TQ = I$ y $Q^T = Q^{-1}$. Así, $S^{-1}AS = \Lambda$ se vuelve especial: es $Q^{-1}AQ = \Lambda$ o $A = Q\Lambda Q^{-1} = Q\Lambda Q^T$. Ahora es posible enunciar uno de los grandes teoremas del álgebra lineal:

5O Una matriz simétrica real puede factorizarse en $A = Q\Lambda Q^T$. Sus vectores característicos ortonormales están en la matriz ortogonal Q, y sus valores característicos están en Λ.

En geometría o mecánica, este es el *teorema de los ejes principales*. Proporciona la elección correcta de ejes para una elipse. Estos ejes son perpendiculares, y apuntan a lo largo de los vectores característicos de la matriz correspondiente. (En la sección 6.2 se relacionan las matrices simétricas con elipses n-dimensionales). En mecánica, los vectores característicos proporcionan las direcciones principales, a lo largo de las cuales hay compresión pura o tensión pura, sin esfuerzo de corte.

En matemáticas, la fórmula $A = Q\Lambda Q^T$ se conoce como *teorema espectral*. Si se multiplican las columnas por los renglones, la matriz A se convierte en una combinación de proyecciones unidimensionales, que son las matrices especiales xx^T de rango 1, multiplicadas por λ:

$$A = Q\Lambda Q^T = \begin{bmatrix} | & & | \\ x_1 & \cdots & x_n \\ | & & | \end{bmatrix}\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}\begin{bmatrix} — x_1^T — \\ \vdots \\ — x_n^T — \end{bmatrix}$$
$$= \lambda_1 x_1 x_1^T + \lambda_2 x_2 x_2^T + \cdots + \lambda_n x_n x_n^T. \quad (10)$$

En el ejemplo los valores característicos son 3 y 1:

Ejemplo 3 $A = \begin{bmatrix}2 & -1\\ -1 & 2\end{bmatrix} = 3\begin{bmatrix}\frac{1}{2} & -\frac{1}{2}\\ -\frac{1}{2} & \frac{1}{2}\end{bmatrix} + \begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2}\end{bmatrix} =$ **combinación de dos proyecciones**

Los vectores característicos, con longitud escalada a 1, son

$$x_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\ -1\end{bmatrix} \quad \text{y} \quad x_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\ 1\end{bmatrix}.$$

Así, las matrices del miembro derecho son $x_1x_1^T$ y $x_2x_2^T$ —columnas por renglones— y son proyecciones sobre la recta que pasa por x_1 y la recta que pasa por x_2.

Todas las matrices simétricas son combinaciones de proyecciones unidimensionales, que son matrices simétricas de rango 1.

Observación Si A es real, y ocurre que sus valores característicos son reales, entonces también sus vectores característicos son reales. Resuelven $(A - \lambda I)x = 0$ y pueden calcularse por eliminación. Sin embargo, no son ortogonales, a menos que A sea simétrica: $A = Q\Lambda Q^T$ lleva a $A^T = A$.

Si A es real, todos los valores característicos complejos se convierten en pares conjugados: $Ax = \lambda x$ y $A\bar{x} = \bar{\lambda}\bar{x}$. *Si $a + ib$ es un valor característico de una matriz real, también lo es $a - ib$.* (Si $A = A^T$, entonces $b = 0$.)

Hablando estrictamente, el teorema espectral $A = Q\Lambda Q^T$ sólo se ha demostrado cuando los valores característicos de A son distintos. Entonces ciertamente hay n vectores característicos independientes y A puede diagonalizarse con seguridad. No obstante, es cierto (consulte la sección 5.6) que *incluso con valores característicos repetidos, una matriz simétrica sigue teniendo un conjunto completo de vectores característicos ortonormales*. El caso extremo es la matriz identidad, que tiene $\lambda = 1$ repetido n veces, y no carece de vectores característicos.

Para terminar el caso complejo se requiere el análogo de una matriz ortogonal real, y el lector puede conjeturar lo que ocurre al requerimiento $Q^TQ = I$. La traspuesta se sustituye por la traspuesta conjugada. La condición se vuelve $U^HU = I$. La nueva letra U refleja la nueva denominación: ***Una matriz compleja con columnas ortonormales se denomina matriz unitaria.***

Matrices unitarias

¿Se nos permite proponer dos analogías? ***Una matriz hermitiana (o simétrica) puede compararse con un número real. Una matriz unitaria (u ortogonal) puede compararse con un número en la circunferencia unitaria:*** un número complejo de valor absoluto 1. Los λs son reales si $A^H = A$, y están en la circunferencia unitaria si $U^HU = I$. Los vectores característicos pueden escalarse a longitud unitaria y hacerse ortonormales.*

Hay dos afirmaciones aún no demostradas para matrices unitarias (incluyendo las ortogonales). En consecuencia, pasamos de inmediato a las tres propiedades de U que corresponden a las propiedades anteriores 1 a 3 de A. Recuerde que las columnas de U son ortonormales:

$$\textbf{Matriz unitaria} \quad U^HU = I, \qquad UU^H = I, \qquad \text{y} \qquad U^H = U^{-1}.$$

Lo anterior lleva directamente a la propiedad 1′, que la multiplicación por U no afecta los productos internos, los ángulos, o las longitudes. La demostración se hace en una línea, justo como lo fue para Q:

Propiedad 1′ $(Ux)^H(Uy) = x^HU^HUy = x^Hy$ y U preserva las longitudes:
La longitud permanece sin cambio $\quad \|Ux\|^2 = x^HU^HUx = \|x\|^2. \quad$ (11)

Propiedad 2′ Todo valor característico de U tiene valor absoluto $|\lambda| = 1$.

Lo anterior se concluye directamente de $Ux = \lambda x$, al comparar las longitudes de los dos miembros: $\|Ux\| = \|x\|$ por la propiedad 1′, y siempre $\|\lambda x\| = |\lambda|\|x\|$. En consecuencia, $|\lambda| = 1$.

Propiedad 3′ Los vectores característicos correspondientes a valores característicos distintos son ortonormales.

*Más tarde se compararán las matrices "antihermitianas" con los números imaginarios puros, y las matrices "normales" con todos los números complejos $a + ib$. Una matriz no normal sin vectores característicos ortogonales no pertenece a ninguna de estas clases, y está fuera de toda analogía.

Se empieza con $Ux = \lambda_1 x$ y $Uy = \lambda_2 y$, y se toman productos internos por la propiedad 1′:

$$x^Hy = (Ux)^H(Uy) = (\lambda_1 x)^H(\lambda_2 y) = \bar{\lambda}_1\lambda_2 x^Hy.$$

Al comparar la izquierda con la derecha, $\bar{\lambda}_1\lambda_2 = 1$ o $x^Hy = 0$. Pero la propiedad 2′ es $\bar{\lambda}_1\lambda_1 = 1$, por lo que no puede tenerse $\bar{\lambda}_1\lambda_2 = 1$. Así, $x^Hy = 0$ y los vectores característicos son ortogonales.

Ejemplo 4 $U = \begin{bmatrix} \cos t & -\operatorname{sen} t \\ \operatorname{sen} t & \cos t \end{bmatrix}$ tiene valores característicos e^{it} y e^{-it}.

Los vectores característicos ortogonales son $x = (1, -i)$ y $y = (1, i)$. (Recuerde tomar conjugados en $x^Hy = 1 + i^2 = 0$). Después de la división entre $\sqrt{2}$ son ortonormales.

A continuación se presenta la *matriz unitaria*, con mucho, la más importante que hay.

Ejemplo 5
$$U = \frac{1}{\sqrt{n}}\begin{bmatrix} 1 & 1 & \cdot & 1 \\ 1 & w & \cdot & w^{n-1} \\ \cdot & \cdot & \cdot & \cdot \\ 1 & w^{n-1} & \cdot & w^{(n-1)^2} \end{bmatrix} = \frac{\text{Matriz de Fourier}}{\sqrt{n}}.$$

El número complejo w está en la circunferencia unitaria al ángulo $\theta = 2\pi/n$. Es igual a $e^{2\pi i/n}$. Sus potencias son equidistantes alrededor de la circunferencia. Esa separación asegura que la suma de todas las n potencias de w —todas las raíces n-ésimas de 1— es cero. Algebraicamente, la suma $1 + w + \cdots + w^{n-1}$ es $(w^n - 1)/(w - 1)$. ¡Y $(w^n - 1)$ es cero!

$$\begin{array}{l}\text{el renglón 1 de } U^H \text{ multiplicado} \\ \text{por la columna 2 de } U \text{ es}\end{array} \quad \frac{1}{n}(1 + w + w^2 + \cdots + w^{n-1}) = \frac{w^n - 1}{w - 1} = 0.$$

$$\begin{array}{l}\text{el renglón } i \text{ de } U^H \text{ multiplicado} \\ \text{por la columna } j \text{ de } U \text{ es}\end{array} \quad \frac{1}{n}(1 + W + W^2 + \cdots + W^{n-1}) = \frac{W^n - 1}{W - 1} = 0.$$

En el segundo caso, $W = w^{j-i}$. Cada elemento de la F original tiene valor absoluto igual a 1. El factor $\sqrt{n}$ encoge las columnas de U en vectores unitarios. ***La identidad fundamental de la transformada finita de Fourier es $U^HU = I$.***

Así, U es una matriz unitaria. Su inversa se ve igual, excepto que w se sustituye por $w^{-1} = e^{-i\theta} = \bar{w}$. Debido a que U es unitaria, su inversa se encuentra trasponiendo (con lo cual todo permanece igual) y conjugando (con lo cual w cambia a $\bar{w}$). La inversa de esta U es $\bar{U}$. Ux puede calcularse rápidamente mediante la ***transformada rápida de Fourier*** según se encontró en la sección 3.5.

Por la propiedad 1′ de las matrices unitarias, la longitud de un vector x es la misma que la longitud de Ux. La energía en el espacio estado es igual a la energía en el espacio transformado. La energía es la suma de $|x_j|^2$, y también es la suma de las energías en las frecuencias por separado. El vector $x = (1, 0, \ldots, 0)$ contiene cantidades iguales de cada componente de la frecuencia, y su transformada de Fourier discreta $Ux = (1, 1, \ldots, 1)/\sqrt{n}$ también tiene longitud 1.

Ejemplo 6
$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Esta es una matriz ortogonal, de modo que por la propiedad 3′ debe tener vectores característicos ortogonales. ¡Estos son las columnas de la matriz de Fourier! El valor absoluto de sus valores característicos debe ser 1.

Hay números $1, w, \ldots, w^{n-1}$ (o $1, i, i^2, i^3$ en este caso de 4 por 4). Es una matriz real, aunque sus valores característicos y vectores característicos son complejos.

Una nota final: Las matrices antihermitianas cumplen $K^H = -K$, así como las matrices simétricas sesgadas satisfacen $K^T = -K$. Sus propiedades se concluyen de inmediato, a partir de su estrecho vínculo con las matrices hermitianas:

Si A es hermitiana, entonces $K = iA$ es anti hermitiana.

Los valores característicos de K son puramente imaginarios, en vez de puramente reales; se multiplica por i. Los vectores característicos no cambian. El ejemplo hermitiano de las páginas previas conduciría a

$$K = iA = \begin{bmatrix} 2i & 3+3i \\ -3+3i & 5i \end{bmatrix} = -K^H.$$

Los elementos diagonales son múltiplos de i (permitiendo el cero). Los valores característicos son $8i$ y $-i$. Los vectores característicos siguen siendo ortogonales, y sigue teniéndose $K = U\Lambda U^H$ —con una U unitaria en vez de una Q ortogonal real, y con $8i$ y $-i$ en la diagonal de Λ.

Esta sección se resume con una tabla de paralelismos entre reales y complejos.

Real contra complejo

$\mathbf{R}^n$ (n componentes reales)	$\leftrightarrow$	$\mathbf{C}^n$ (n componentes complejas)
longitud: $\|x\|^2 = x_1^2 + \cdots + x_n^2$	$\leftrightarrow$	longitud: $\|x\|^2 = \lvert x_1\rvert^2 + \cdots + \lvert x_n\rvert^2$
traspuesta: $A_{ij}^T = A_{ji}$	$\leftrightarrow$	traspuesta hermitiana: $A_{ij}^H = \overline{A_{ji}}$
$(AB)^T = B^TA^T$	$\leftrightarrow$	$(AB)^H = B^HA^H$
producto interno: $x^Ty = x_1y_1 + \cdots + x_ny_n$	$\leftrightarrow$	$x^Hy = \bar{x}_1y_1 + \cdots + \bar{x}_ny_n$
$(Ax)^Ty = x^T(A^Ty)$	$\leftrightarrow$	$(Ax)^Hy = x^H(A^Hy)$
ortogonalidad: $x^Ty = 0$	$\leftrightarrow$	ortogonalidad: $x^Hy = 0$
matrices simétricas: $A^T = A$	$\leftrightarrow$	matrices hermitianas: $A^H = A$
$A = Q\Lambda Q^{-1} = Q\Lambda Q^T$ (Λ real)	$\leftrightarrow$	$A = U\Lambda U^{-1} = U\Lambda U^H$ (Λ real)
simétrica sesgada $K^T = -K$	$\leftrightarrow$	antihermitiana $K^H = -K$
ortogonal $Q^TQ = I$ o bien, $Q^T = Q^{-1}$	$\leftrightarrow$	$U^HU = I$ unitaria o bien $U^H = U^{-1}$
$(Qx)^T(Qy) = x^Ty$ y $\|Qx\| = \|x\|$	$\leftrightarrow$	$(Ux)^H(Uy) = x^Hy$ y $\|Ux\| = \|x\|$

Las columnas, los renglones, y los vectores característicos de Q y U son ortonormales, y todo $\lvert\lambda\rvert = 1$.

Conjunto de problemas 5.5

1. Para los números complejos $3 + 4i$ y $1 - i$:
   a) Encuentre sus posiciones en el plano complejo.
   b) Encuentre su suma y su producto.
   c) Encuentre sus conjugados y sus valores absolutos.

   ¿Los números originales están dentro o fuera de la circunferencia unitaria?

2. ¿Qué puede decir sobre
   a) la suma de un número complejo y su conjugado?
   b) el conjugado de un número que está sobre la circunferencia unitaria?
   c) el producto de dos números que están sobre la circunferencia unitaria?
   d) la suma de dos números que están sobre la circunferencia unitaria?

3. Si $x = 2 + i$ y $y = 1 + 3i$, encuentre $\bar{x}$, $x\bar{x}$, xy, $1/x$ y x/y. Compruebe que el valor absoluto $\lvert xy\rvert$ es igual a $\lvert x\rvert$ multiplicado por $\lvert y\rvert$, y que el valor absoluto $\lvert 1/x\rvert$ es igual a 1 dividido entre $\lvert x\rvert$.

4. Encuentre a y b para los números complejos $a + ib$ a los ángulos $\theta = 30°, 60°, 90°$ sobre la circunferencia unitaria. Compruebe por multiplicación directa que el cuadrado del primero es el segundo, y que el cubo del primero es el tercero.

5. *a*) Si $x = re^{i\theta}$, ¿cuáles son x^2, x^{-1}, y $\bar{x}$ en coordenadas polares? ¿Dónde están los números complejos que tienen $x^{-1} = \bar{x}$?

   b) En $t = 0$, el número complejo $e^{(-1+i)t}$ es igual a la unidad. Trace su trayectoria en el plano complejo cuando t crece desde 0 hasta 2π.

6. Encuentre las longitudes y el producto interno de

$$x = \begin{bmatrix} 2-4i \\ 4i \end{bmatrix} \quad \text{y} \quad y = \begin{bmatrix} 2+4i \\ 4i \end{bmatrix}.$$

7. Escriba la matriz A^H, y calcule $C = A^HA$ si

$$A = \begin{bmatrix} 1 & i & 0 \\ i & 0 & 1 \end{bmatrix}.$$

   ¿Cuál es la relación entre C y C^H? ¿Esto se cumple siempre que C se construye a partir de alguna A^HA?

8. *a*) Con la A precedente, use eliminación para resolver $Ax = 0$.
   b) Demuestre que el espacio nulo que calculó es ortogonal a $C(A^H)$ *y no al espacio renglón de costumbre* $C(A^T)$. Los cuatro espacios fundamentales en el caso complejo son $N(A)$ y $C(A)$ como antes, y luego $N(A^H)$ y $C(A^H)$.

9. *a*) ¿Cómo está relacionado el determinante de A^H con el determinante de A?
   b) Demuestre que el determinante de cualquier matriz hermitiana es real.

10. *a*) ¿Cuántos grados de libertad hay en una matriz simétrica real, en una matriz diagonal real, y en una matriz ortogonal real? (La primera respuesta es la suma de las otras dos, ya que $A = Q\Lambda Q^T$.)
    b) Demuestre que las matrices hermitianas A de 3 por 3 y también la U unitaria tienen 9 grados de libertad (las columnas de U pueden multiplicarse por cualquier $e^{i\theta}$).

11. Escriba P, Q y R en la forma $\lambda_1x_1x_1^H + \lambda_2x_2x_2^H$ del teorema espectral:

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \quad Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix}.$$

12. Proporcione una razón si es verdadero o un contraejemplo si es falso:

a) Si A es hermitiana, entonces $A + iI$ es invertible.
b) Si Q es ortogonal, entonces $Q + \frac{1}{2}I$ es invertible.
c) Si A es real, entonces $A + iI$ es invertible.

13. Suponga que A es una matriz simétrica de 3 por 3 con valores característicos 0, 1, 2.

a) ¿Qué propiedades pueden garantizarse para los vectores característicos unitarios correspondientes u, v, w?
b) En términos de u, v, w, describa el espacio nulo, el espacio nulo izquierdo, el espacio renglón, y el espacio columna de A.
c) Encuentre un vector x que cumpla $Ax = v + w$, ¿x es único?
d) ¿En qué condiciones sobre b, $Ax = b$ tiene una solución?
e) Si u, v, w son las columnas de S, ¿cuáles son S^{-1} y $S^{-1}AS$?

14. En la siguiente lista, ¿qué clases de matrices contienen a A, y cuáles contienen a B?

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{y} \quad B = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Ortogonales, invertibles, proyección, permutación, hermitianas, de rango 1, *diagonalizables, de Markov*. Encuentre los valores característicos de A y B.

15. ¿Cuál es la dimensión del espacio S de todas las matrices simétricas reales de n por n? El teorema espectral establece que toda matriz simétrica es una combinación de n matrices proyección. Debido a que la dimensión excede a n, ¿cómo se explica esta diferencia?

16. Escriba un hecho importante sobre los valores característicos de cada uno de los siguientes incisos.

a) Una matriz simétrica real.
b) Una matriz estable: todas las soluciones de $du/dt = Au$ tienden a cero.
c) Una matriz ortogonal.
d) Una matriz de Markov.
e) Una matriz defectuosa (no diagonalizable).
f) Una matriz singular.

17. Demuestre que si U y V son unitarias, también lo es $U\,V$. Use el criterio $U^{\mathrm{H}}U = I$.

18. Demuestre que una matriz unitaria tiene $|\det U| = 1$, aunque quizá $\det U$ es diferente de $\det U^{\mathrm{H}}$. Describa todas las matrices de 2 por 2 que son unitarias.

19. Encuentre una tercera columna, de modo que U sea unitaria. ¿Cuánta libertad hay en la columna 3?

$$U = \begin{bmatrix} 1/\sqrt{3} & i/\sqrt{2} & \\ 1/\sqrt{3} & 0 & \\ i/\sqrt{3} & 1/\sqrt{2} & \end{bmatrix}.$$

20. Diagonalice la matriz antihermitiana de 2 por 2 $K = \begin{bmatrix} i & i \\ i & i \end{bmatrix}$, cuyos elementos son todos $\sqrt{-1}$.

Calcule $e^{Kt} = Se^{\Lambda t}S^{-1}$ y compruebe que e^{Kt} es unitaria. ¿Cuál es la derivada de e^{Kt} en $t = 0$?

21. Describa todas las matrices de 3 por 3 que simultáneamente son hermitianas, unitarias, y diagonales. ¿Cuántas hay?

22. Toda matriz Z puede separarse en una parte hermitiana y una parte antihermitiana, $Z = A + K$, así como un número complejo z puede separarse en $a + ib$. La parte real de z es la mitad de $z + \bar{z}$, y la "parte real" de Z es la mitad de $Z + Z^{\mathrm{H}}$. Encuentre una fórmula semejante para la "parte imaginaria" K, y separe estas matrices en $A + K$:

$$Z = \begin{bmatrix} 3+i & 4+2i \\ 0 & 5 \end{bmatrix} \quad \text{y} \quad Z = \begin{bmatrix} i & i \\ -i & i \end{bmatrix}.$$

23. Demuestre que las columnas de la matriz de Fourier de 4 por 4 en el ejemplo 5 son vectores característicos de la matriz permutación P, en el ejemplo 6.

24. Para la permutación en el ejemplo 6, escriba la *matriz circulante* $C = c_0I + c_1P + c_2P^2 + c_3P^3$. (Su matriz vector característico es nuevamente una matriz de Fourier.) También escriba las cuatro componentes del producto matriz-vector Cx, que es la *convolución* de $c = (c_0, c_1, c_2, c_3)$ y $x = (x_0, x_1, x_2, x_3)$.

25. Para una matriz circulante $C = F\Lambda F^{-1}$, ¿por qué es más rápido multiplicar por F^{-1}, luego por Λ, y luego por F (regla de convolución), que multiplicar directamente por C?

26. Encuentre las longitudes de $u = (1 + i, 1 - i, 1 + 2i)$ y $v = (i, i, i)$. También encuentre $u^{\mathrm{H}}v$ y $v^{\mathrm{H}}u$.

27. Demuestre que $A^{\mathrm{H}}A$ siempre es una matriz hermitiana. Calcule $A^{\mathrm{H}}A$ y AA^{H}:

$$A = \begin{bmatrix} i & 1 & i \\ 1 & i & i \end{bmatrix}.$$

28. Si $Az = 0$, entonces $A^{\mathrm{H}}Az = 0$. Si $A^{\mathrm{H}}Az = 0$, multiplique por z^{H} para demostrar que $Az = 0$. Los espacios nulos de A y $A^{\mathrm{H}}A$ son ______. $A^{\mathrm{H}}A$ es una matriz hermitiana invertible cuando el espacio nulo de A sólo contiene a $z =$ ______.

29. Cuando una matriz hermitiana se multiplica por un número real c, ¿cA sigue siendo hermitiana? Si $c = i$, demuestre que iA es antihermitiana. Las matrices hermitianas de 3 por 3 constituyen un subespacio, en el supuesto de que los "escalares" sean números reales.

30. ¿Qué clases de matrices P pertenecen a: ortogonales, invertibles, hermitianas, unitarias, factorizables en LU, factorizables en QR?

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

31. Calcule P^2, P^3, y P^{100} en el problema 30. ¿Cuáles son los valores característicos de P?

32. Encuentre los vectores característicos unitarios de P en el problema 30, y luego escríbalos en las columnas de una matriz unitaria U. ¿Qué propiedad de P hace ortogonales a estos vectores característicos?

33. Escriba la *matriz circulante* de 3 por 3 $C = 2I + 5P + 4P^2$. Tiene los mismos vectores característicos que P en el problema 30. Encuentre sus valores característicos.

34. Si U es unitaria y Q es una matriz ortogonal real, demuestre que U^{-1} es unitaria, y también que UQ es unitaria. Empiece con $U^H U = I$ y $Q^T Q = I$.

35. Diagonalice A (λs reales) y K (λs imaginarios) para llegar a $U\Lambda U^H$:

$$A = \begin{bmatrix} 0 & 1-i \\ i+1 & 1 \end{bmatrix} \qquad K = \begin{bmatrix} 0 & -1+i \\ 1+i & i \end{bmatrix}$$

36. Diagonalice la siguiente matriz ortogonal para llegar a $Q = U\Lambda U^H$. Ahora todos los λs son ______.

$$Q = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}.$$

37. Diagonalice la siguiente matriz unitaria, V para llegar a $V = U\Lambda U^H$. Nuevamente, todos los $|\lambda| = 1$:

$$V = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1-i \\ 1+i & -1 \end{bmatrix}.$$

38. Si $v_1, \ldots, v_n$ es una base ortonormal de $\mathbf{C}^n$, la matriz con estas columnas es una matriz ______. Demuestre que cualquier vector z es igual a $(v_1^H z)v_1 + \cdots + (v_n^H z)v_n$.

39. Las funciones e^{-ix} y e^{ix} son ortogonales en el intervalo $0 \le x \le 2\pi$ porque su producto interno *complejo* es $\int_0^{2\pi}$ ______ $= 0$.

40. Los vectores $v = (1, i, 1)$, $w = (i, 1, 0)$ y $z =$ ______ son una base ortogonal de ______.

41. Si $A = R + iS$ es una matriz hermitiana, ¿las matrices R y S, son simétricas?

42. La dimensión (compleja) de $\mathbf{C}^n$ es ______. Encuentre una base no real de $\mathbf{C}^n$.

43. Describa todas las matrices de 1 por 1 que sean hermitianas y también unitarias. Haga lo mismo para las matrices de 2 por 2.

44. ¿Cómo están relacionados los valores característicos de A^H (una matriz cuadrada) con los valores característicos de A?

45. Si $u^H u = 1$, demuestre que $I - 2uu^H$ es hermitiana y también unitaria. La matriz uu^H de rango 1 es la proyección sobre ¿qué recta en $\mathbf{C}^n$?

46. Si $A + iB$ es una matriz unitaria (A y B son reales), demuestre que $Q = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}$ es una matriz ortogonal.

47. Si $A + iB$ es una matriz hermitiana (A y B son reales), demuestre que $\begin{bmatrix} A & -B \\ B & A \end{bmatrix}$ es simétrica.

48. Demuestre que la inversa de una matriz hermitiana también es hermitiana.

49. Diagonalice la siguiente matriz, construyendo su matriz valor característico Λ y su matriz vector característico S:

$$A = \begin{bmatrix} 2 & 1-i \\ 1+i & 3 \end{bmatrix} = A^H.$$

50. Una matriz con vectores característicos ortonormales es de la forma $A = U\Lambda U^{-1} = U\Lambda U^H$. *Demuestre que* $AA^H = A^H A$. Éstas son exactamente las *matrices normales*.

5.6 TRANSFORMACIONES DE SEMEJANZA

Virtualmente, cada paso de este capítulo implicó la combinación $S^{-1}AS$. Los vectores característicos de A se fueron en las columnas S, con lo cual S^1AS se volvió una matriz diagonal (denominada Λ). Una vez que A era simétrica, en vez de S, se escribió Q, escogiendo que los vectores característicos fuesen ortonormales. En el caso complejo, cuando A es hermitiana, se escribió U, que sigue siendo la matriz de vectores característicos. Ahora se consideran todas las combinaciones $M^{-1}AM$, *formadas con cualquier M invertible en la derecha y su inversa en la izquierda.* La matriz vector característico invertible S podría no existir (el caso defectuoso), o podría no ser conocida, e incluso podríamos no querer utilizarla.

Primero un nuevo comentario: ***Las matrices A y $M^{-1}AM$ son "semejantes".*** Pasar de una a otra es una ***transformación de semejanza***. Es el paso natural para las ecuaciones diferenciales o matrices de potencias o valores característicos; así como los pasos de la eliminación eran naturales para $Ax = b$. La eliminación multiplicaba A por la izquierda por L^{-1}, pero no lo hacía por la derecha por L. De modo que U no es semejante a A, y los pivotes *no son* los valores característicos.

Toda una familia de matrices $M^{-1}AM$ es semejante a A, y hay dos preguntas:

1. ¿Qué tienen en común estas matrices semejantes $M^{-1}AM$?
2. Con una elección especial de M, ¿qué forma especial puede obtenerse mediante $M^{-1}AM$?

La respuesta final la proporciona la ***forma de Jordan***, con la que termina el capítulo.

Estas combinaciones $M^{-1}AM$ se presentan en una ecuación diferencial o en diferencias, cuando un "cambio de variables" $u = Mv$ introduce la nueva incógnita v:

$$\frac{du}{dt} = Au \quad \text{se convierte en} \quad M\frac{dv}{dt} = AMv, \quad \text{o bien,} \quad \frac{dv}{dt} = M^{-1}AMv$$

$$u_{n+1} = Au_n \quad \text{se convierte en} \quad Mv_{n+1} = AMv_n, \quad \text{o bien,} \quad v_{n+1} = M^{-1}AMv_n.$$

La nueva matriz en la ecuación es $M^{-1}AM$. En el caso especial $M = S$, el sistema no está acoplado porque $\Lambda = S^{-1}AS$ es diagonal. Los vectores característicos evolucionan de manera independiente. Esta es la simplificación máxima, aunque también son de utilidad otras Ms. Se intentará que trabajar con $M^{-1}AM$ sea más fácil que hacerlo con A.

La familia de matrices $M^{-1}AM$ incluye a A misma, eligiendo $M = I$. Cualquiera de estas matrices semejantes puede aparecer en las ecuaciones diferencial y en diferencias, mediante el cambio $u = Mv$, por lo que deben tener algo en común, como es el caso: ***las matrices semejantes comparten los mismos valores característicos.***

> **5P** Suponga que $B = M^{-1}AM$. Entonces A y B tienen los **mismos valores característicos. Todo vector característico x de A corresponde a un vector característico $M^{-1}x$ de B.**

Empiece con $Ax = \lambda x$ y sustituya $A = MBM^{-1}$:

$$\textbf{Mismo valor característico} \quad MBM^{-1}x = \lambda x \quad \text{que es} \quad B(M^{-1}x) = \lambda(M^{-1}x). \qquad (1)$$

El valor característico de B sigue siendo λ. El vector característico ha cambiado de x a $M^{-1}x$.

También puede comprobarse que $A - \lambda I$ y $B - \lambda I$ tienen el mismo determinante:

Producto de matrices $\quad B - \lambda I = M^{-1}AM - \lambda I = M^{-1}(A - \lambda I)M$

Regla del producto $\quad \det(B - \lambda I) = \det M^{-1} \det(A - \lambda I) \det M = \det(A - \lambda I).$

Los polinomios $\det(A - \lambda I)$ y $\det(B - \lambda I)$ son iguales. Sus raíces —los valores característicos de A y B— son los mismos. A continuación se presentan las matrices B que son semejantes a A.

Ejemplo 1 $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ tiene valores característicos 1 y 0. Cada B es $M^{-1}AM$:

$$\text{Si } M = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}, \text{ entonces } B = \begin{bmatrix} 1 & b \\ 0 & 0 \end{bmatrix}: \text{ triangular con } \lambda = 1 \text{ y } 0.$$

$$\text{Si } M = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \text{ entonces } B = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}: \text{ proyección con } \lambda = 1 \text{ y } 0.$$

$$\text{Si } M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \text{ entonces } B = \text{ una matriz arbitraria con } \lambda = 1 \text{ y } 0.$$

En este caso es posible producir cualquier B con los valores característicos correctos. Se trata de un caso fácil, ya que los valores característicos 1 y 0 son distintos. La matriz diagonal A en realidad era Λ, el elemento distinguido de esta familia de matrices semejantes. La forma de Jordan estará preocupada sobre valores característicos repetidos y una posible carencia de vectores característicos. Todo lo que se dice ahora es que cada $M^{-1}AM$ tiene el mismo número de vectores característicos independientes que A (cada vector característico se multiplica por M^{-1}).

El primer paso es considerar las transformaciones lineales que están detrás de las matrices. Rotaciones, reflexiones, y proyecciones actúan sobre el espacio n-dimensional. La transformación puede ocurrir sin álgebra lineal, aunque ésta resulta en multiplicación de matrices.

Cambio de base = Transformación de semejanza

La matriz semejante $B = M^{-1}AM$ está estrechamente relacionada con A, si se regresa al estudio de las transformaciones lineales. Recuerde la idea clave: ***Toda transformación lineal está representada por una matriz.*** ¡La matriz depende de la elección de la base! *Si la base se cambia por M, entonces la matriz A se cambia por una matriz semejante B.*

Las matrices semejantes representan la misma transformación T respecto a bases diferentes. El álgebra es casi directa: Suponga que se tiene una base $v_1, \ldots, v_n$. La j-ésima columna de A se obtiene al aplicar T a v_j:

$$Tv_j = \text{combinación de los vectores de la base} = a_{1j}v_1 + \cdots + a_{nj}v_n. \tag{2}$$

Para una nueva base $V_1, \ldots, V_n$, la nueva matriz B se construye de la misma forma: TV_j = combinación de los Vs $= b_{1j}V_1 + \cdots + b_{nj}V_n$. Pero también cada V debe ser una combinación de los vectores de la base anterior: los $V_j = \sum m_{ij}v_i$. Esta matriz M en realidad representa la *transformación identidad* (!) cuando todo lo que ocurre es el cambio de base (T es I). La matriz inversa M^{-1} también representa la transformación identidad, cuando la base se cambia de los vs de regreso a los Vs. Así, la regla del producto proporciona el resultado que se busca:

5Q Las matrices A y B que representan la misma transformación lineal T respecto a dos bases diferentes (los vs y los Vs) son ***semejantes:***

$$\begin{matrix} [T]_{V \text{ a } V} & = [I]_{v \text{ a } V} & [T]_{v \text{ a } v} & [I]_{V \text{ a } v} \\ B & = M^{-1} & A & M. \end{matrix} \tag{3}$$

Considero que la mejor forma de explicar $B = M^{-1}AM$ es con un ejemplo. Suponga que T es *la proyección sobre la recta L* al ángulo θ. Esta transformación lineal es descrita completamente sin ayuda de una base. Sin embargo, para representar T con una matriz se requiere de una base. En la figura 5.5 se presentan dos posibilidades, la base estándar $v_1 = (1, 0)$, $v_2 = (0, 1)$ y una base V_1, V_2 escogida especialmente para T.

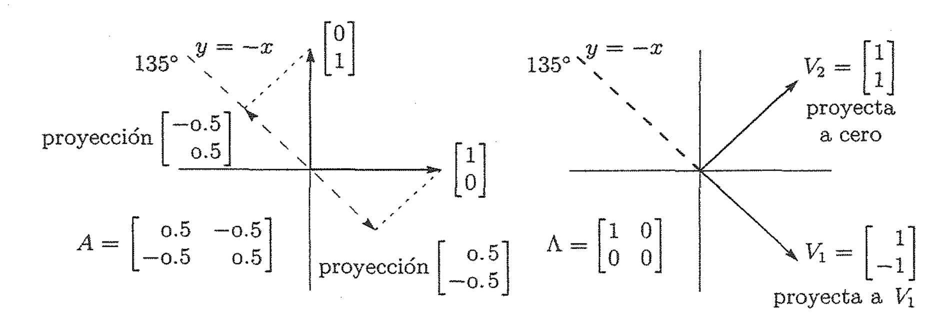


Figura 5.5 Cambio de base, con la finalidad de hacer diagonal a la matriz proyección.

De hecho, $TV_1 = V_1$ (porque V_1 ya está en la recta L) y $TV_2 = 0$ (porque V_2 es perpendicular a la recta). En esa base de vectores característicos, la matriz es diagonal:

$$\textbf{Base de vectores característicos} \qquad B = [T]_{V \text{ a } V} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

La otra cuestión es el cambio de matriz base M. Para ello, V_1 se expresa como una combinación $v_1 \cos\theta + v_2 \operatorname{sen}\theta$ y estos coeficientes se escriben en la columna 1. De manera semejante, V_2 (o IV_2, la transformación es la identidad) es $-v_1 \operatorname{sen}\theta + v_2 \cos\theta$, con lo que se obtiene la columna 2:

$$\textbf{Cambio de base} \qquad M = [I]_{V \text{ a } v} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}.$$

La matriz inversa M^{-1} (que aquí es la traspuesta) va de v a V. Combinada con B y M, proporciona la matriz proyección en la base estándar de vs:

$$\textbf{Base estándar} \qquad A = MBM^{-1} = \begin{bmatrix} c^2 & cs \\ cs & s^2 \end{bmatrix}.$$

Es posible resumir la cuestión importante. La manera de simplificar la matriz A —de hecho diagonalizarla— significa encontrar sus vectores característicos. Éstos van en las columnas de M (o de S) y $M^{-1}AM$ es diagonal. El algebrista afirma lo mismo en el lenguaje de las transformaciones lineales: *escoger una base que conste de vectores característicos.* La base estándar llevó a A, lo cual no era sencillo. La base correcta condujo a B, que era diagonal.

Nuevamente se recalca que $M^{-1}AM$ no surge cuando se resuelve $Ax = b$. Ahí la operación básica fue multiplicar A (¡sólo por el lado izquierdo!) por una matriz que resta un múltiplo de un renglón de otro. Esta transformación preserva el espacio nulo y el espacio renglón de A: normalmente cambia los valores característicos.

En realidad, los valores característicos se calculan por medio de una sucesión de semejanzas. La matriz avanza poco a poco hacia una forma triangular, y los valores característicos aparecen de manera gradual sobre la diagonal principal. (Esta sucesión se describe

en el capítulo 7). Eso es mucho mejor que tratar de calcular det $(A - \lambda I)$, cuyas raíces deben ser los valores característicos. Para una gran matriz, numéricamente es imposible concentrar toda esta información en el polinomio y obtenerla de nuevo.

Formas triangulares con una M unitaria

Nuestro primer movimiento más allá de la matriz vector característico $M = S$ es un poco ilógico: en vez de una M más general, se avanza en dirección opuesta y *M se restringe de modo que sea unitaria*. Con esta restricción, $M^{-1}AM$ puede alcanzar una forma T triangular. Las columnas de $M = U$ son ortonormales (en el caso real podría escribirse $M = Q$). A menos que los vectores característicos de A sean ortogonales, una diagonal $U^{-1}AU$ es imposible. Sin embargo, el "lema de Schur" en **5R** es muy útil, por lo menos para la teoría. (El resto de este capítulo está dedicado más a la teoría que a aplicaciones. La forma de Jordan es independiente de esta forma triangular).

5R Hay una matriz unitaria $M = U$ tal que $U^{-1}AU = T$ es triangular. Los valores característicos de A aparecen a lo largo de la diagonal de esta matriz semejante T.

Demostración Toda matriz, por ejemplo de 4 por 4, tiene por lo menos un valor característico λ_1. En el peor de los casos, puede estar repetido cuatro veces. En consecuencia, A tiene por lo menos un vector característico unitario x_1, que se coloca en la *primera columna de U*. En esta etapa, es imposible determinar las otras tres columnas, por lo que la matriz se completa de cualquier forma que la deje unitaria, y se la denomina U_1. (El proceso de Gram-Schmidt garantiza que esto es posible). $Ax_1 = \lambda_1 x_1$ en la columna 1 significa que el producto $U_1^{-1}AU_1$ empieza de forma correcta:

$$AU_1 = U_1 \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix} \quad \text{conduce a} \quad U_1^{-1}AU_1 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix}.$$

Luego, se trabaja con la submatriz de 3 por 3 en la esquina inferior derecha. Tiene un vector característico unitario x_2, que se vuelve la primera columna de una matriz unitaria M_2:

$$\text{Si } U_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & M_2 & \\ 0 & & & \end{bmatrix} \quad \text{entonces} \quad U_2^{-1}(U_1^{-1}AU_1)U_2 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{bmatrix}.$$

En el último paso, un vector característico de la matriz de 2 por 2 en la esquina inferior derecha pasa a una M_3 unitaria, que se coloca en la esquina de U_3:

$$\textbf{Triangular} \quad U_3^{-1}(U_2^{-1}U_1^{-1}AU_1U_2)U_3 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & \lambda_3 & * \\ 0 & 0 & 0 & * \end{bmatrix} = T.$$

El producto $U = U_1U_2U_3$ sigue siendo una matriz unitaria, y $U^{-1}AU = T$. ■

Este lema es válido para todas las matrices, sin la hipótesis de que A es diagonalizable. Puede usarse para demostrar que ***las potencias de A^k tienden a cero cuando todos los $|\lambda_i| < 1$, y las exponenciales e^{At} tienden a cero cuando todos los* Re $\lambda_i < 0$**, incluso sin el conjunto completo de vectores característicos que se supuso en las secciones 5.3 y 5.4.

Ejemplo 2 $A = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$ tiene el valor característico $\lambda = 1$ (dos veces).

La única línea de vectores característicos pasa por (1, 1). Después de dividir entre $\sqrt{2}$, ésta es la primera columna de U, y la $U^{-1}AU = T$ triangular tiene los valores característicos sobre su diagonal:

$$U^{-1}AU = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = T. \quad (4)$$

Diagonalización de matrices simétricas y hermitianas

Esta forma triangular mostrará que cualquier matriz simétrica o hermitiana —sin importar que sus valores característicos sean *distintos o no*— tiene un conjunto completo de vectores característicos ortonormales. Se requiere una matriz unitaria tal que $U^{-1}AU$ sea *diagonal*. El lema de Schur justamente acaba de encontrarla. Esta T triangular debe ser diagonal, ya que también es hermitiana cuando $A = A^H$:

$$T = T^H \qquad (U^{-1}AU)^H = U^H A^H (U^{-1})^H = U^{-1}AU.$$

La matriz diagonal $U^{-1}AU$ representa un teorema clave en álgebra lineal.

5S (Teorema espectral) Toda A simétrica real puede diagonalizarse por una matriz ortogonal Q. Toda matriz hermitiana puede diagonalizarse por una matriz unitaria U.

$$\textbf{(real)} \quad Q^{-1}AQ = \Lambda \quad \text{o bien,} \quad A = Q\Lambda Q^T$$

$$\textbf{(compleja)} \quad U^{-1}AU = \Lambda \quad \text{o bien,} \quad A = U\Lambda U^H$$

Las columnas de Q (o de U) contienen vectores característicos ortonormales de A.

Observación 1 En el caso simétrico real, los valores característicos y los vectores característicos son reales en cada paso. Así se obtiene una U unitaria *real*, que es una matriz ortogonal.

Observación 2 A es el límite de matrices simétricas con valores característicos *distintos*. A medida que se tiende al límite, los vectores característicos permanecen perpendiculares. Esto puede fallar si $A \neq A^T$:

$$A(\theta) = \begin{bmatrix} 0 & \cos\theta \\ 0 & \operatorname{sen}\theta \end{bmatrix} \quad \text{tiene vectores característicos} \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} \cos\theta \\ \operatorname{sen}\theta \end{bmatrix}.$$

Cuando $\theta \to 0$, el *único* vector característico de la matriz no diagonalizable $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ es $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Ejemplo 3 El teorema espectral establece que esta $A = A^T$ puede diagonalizarse:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{con valores característicos repetidos } \lambda_1 = \lambda_2 = 1 \text{ y } \lambda_3 = -1.$$

$\lambda = 1$ tiene un plano de vectores característicos, y se escoge un par x_1 y x_2:

$$x_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\end{bmatrix} \quad \text{y} \quad x_2 = \begin{bmatrix}0\\0\\1\end{bmatrix}; \quad \text{y} \quad x_3 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\\0\end{bmatrix} \text{ para } \lambda_3 = -1.$$

Estas son las columnas de Q. Al separar $A = Q\Lambda Q^{\mathrm{T}}$ en 3 columnas multiplicadas por 3 renglones se obtiene

$$A = \begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix} = \lambda_1\begin{bmatrix}\frac{1}{2}&\frac{1}{2}&0\\\frac{1}{2}&\frac{1}{2}&0\\0&0&0\end{bmatrix} + \lambda_2\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix} + \lambda_3\begin{bmatrix}\frac{1}{2}&-\frac{1}{2}&0\\-\frac{1}{2}&\frac{1}{2}&0\\0&0&0\end{bmatrix}.$$

Debido a que $\lambda_1 = \lambda_2$, estas dos primeras proyecciones $x_1x_1^{\mathrm{T}}$ y $x_2x_2^{\mathrm{T}}$ (cada una de rango 1) se combinan para proporcionar una proyección P_1 de rango 2 (sobre el plano de vectores característicos). Así, A es

$$\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix} = \lambda_1 P_1 + \lambda_3 P_3 = (+1)\begin{bmatrix}\frac{1}{2}&\frac{1}{2}&0\\\frac{1}{2}&\frac{1}{2}&0\\0&0&1\end{bmatrix} + (-1)\begin{bmatrix}\frac{1}{2}&-\frac{1}{2}&0\\-\frac{1}{2}&\frac{1}{2}&0\\0&0&0\end{bmatrix}. \qquad (5)$$

Toda matriz hermitiana con k valores característicos distintos tiene una ***descomposición espectral*** *en* $A = \lambda_1 P_1 + \cdots + \lambda_k P_k$, *donde* P_i *es la proyección sobre el espacio característico para* λ_i. Debido a que hay un conjunto completo de vectores característicos, la suma de las proyecciones es igual a la identidad. Y como los espacios característicos son ortogonales, dos proyecciones producen cero: $P_j P_i = 0$.

Se está muy cerca de responder una pregunta importante, por lo que se continúa: **¿Para qué matrices se cumple que $T = \Lambda$?** ¡Todas las Ts simétricas, simétricas sesgadas y ortogonales son diagonales! Las matrices hermitianas, antihermitianas y unitarias, también están en esta clase. Corresponden a números en el *eje real*, el *eje imaginario*, y en la *circunferencia unitaria*. Ahora se busca toda la clase, que corresponde a todos los números complejos. Las matrices se denominan "normales".

5T La matriz N es ***normal*** si conmuta con N^{H}: $NN^{\mathrm{H}} = N^{\mathrm{H}}N$. Para tales matrices, y no otras, la triangular $T = U^{-1}NU$ es la diagonal Λ. Las matrices normales son exactamente aquellas que tienen un ***conjunto completo de vectores característicos ortonormales***.

Ciertamente, las matrices simétricas y las matrices hermitianas son normales: Si $A = A^{\mathrm{H}}$, entonces ambas AA^{H} y $A^{\mathrm{H}}A$ son iguales a A^2. Las matrices ortogonales y las unitarias también son normales: ambas UU^{H} y $U^{\mathrm{H}}U$ son iguales a I. Para cualquier matriz normal bastan dos pasos:

1. Si N es normal, también lo es la triangular $T = U^{-1}NU$:

$$TT^{\mathrm{H}} = U^{-1}NUU^{\mathrm{H}}N^{\mathrm{H}}U = U^{-1}NN^{\mathrm{H}}U = U^{-1}N^{\mathrm{H}}NU = U^{\mathrm{H}}N^{\mathrm{H}}UU^{-1}NU = T^{\mathrm{H}}T.$$

2. ¡Una T triangular que sea normal debe ser diagonal! (Consulte los problemas 19 y 20 al final de esta sección).

Entonces, si N es normal, la triangular $T = U^{-1}NU$ debe ser diagonal. Debido a que T tiene los mismos valores característicos que N, debe ser Λ. Los vectores característicos de N son las columnas de U, y son ortonormales. Éste es el buen caso. A continuación se regresa de las mejores matrices (*normales*) a las peores posibles (*defectuosas*).

$$\textbf{Normal} \quad N = \begin{bmatrix}2&1\\-1&2\end{bmatrix} \qquad\qquad \textbf{Defectuosa} \quad A = \begin{bmatrix}2&1\\0&2\end{bmatrix}.$$

La forma de Jordan

Esta sección ha hecho su mejor esfuerzo a la vez que demandó que M sea una matriz unitaria U. Se obtuvo $M^{-1}AM$ en una forma triangular T. Ahora se retira esta restricción sobre M. Se permite cualquier matriz, y el objetivo es hacer $M^{-1}AM$ *lo más diagonal posible*.

El resultado de este esfuerzo supremo para diagonalización es la ***forma de Jordan J***. Si A tiene un conjunto completo de vectores característicos, se toma $M = S$ y se llega a $J = S^{-1}AS = \Lambda$. Así, la forma de Jordan coincide con la diagonal Λ. Esto es imposible para una matriz defectuosa (no diagonalizable). *Para todo vector característico faltante, la forma de Jordan tiene un* 1 *justo arriba de su diagonal principal*. Los valores característicos aparecen sobre la diagonal porque J es triangular. Y siempre es posible desacoplar valores característicos distintos.

Lo único que puede (o no) requerir un 1 fuera de la diagonal en J es un λ repetido.

5U Si A tiene s vectores característicos independientes, es semejante a una matriz con s bloques:

$$\textbf{Forma de Jordan} \quad J = M^{-1}AM = \begin{bmatrix}J_1&&&\\&\cdot&&\\&&\cdot&\\&&&J_s\end{bmatrix}. \qquad (6)$$

Cada bloque de Jordan J_i es una matriz triangular que sólo tiene un valor característico λ_i y un solo vector característico:

$$\textbf{Bloque de Jordan} \quad J_i = \begin{bmatrix}\lambda_i&1&&\\&\lambda_i&\cdot&\\&&\cdot&1\\&&&\lambda_i\end{bmatrix}. \qquad (7)$$

El mismo λ_i aparece en varios bloques, si tiene varios vectores característicos independientes. Dos matrices son semejantes si y sólo si comparten la misma forma de Jordan J.

Muchos autores han hecho de este teorema el punto clímax de su curso de álgebra lineal. Francamente, considero que esto es un error. Ciertamente es verdad que no todas las matrices son diagonalizables, y la forma de Jordan es el caso más general. Simplemente por ello, su construcción es técnica y extremadamente inestable. (Un ligero cambio en A puede regresar todos los vectores característicos faltantes, y eliminar los 1s fuera de la diagonal.) En consecuencia, el lugar correcto para tratar los detalles es el apéndice, y la mejor manera de comenzar a trabajar con la forma de Jordan es considerando algunos ejemplos específicos y manipulables.

Ejemplo 4 $T = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ y $A = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ conducen a $J = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

Estas cuatro matrices tienen valores característicos 1 y 1 con sólo *un vector característico*, de modo que J consta de *un bloque*. A continuación se comprobará este hecho. Todos los determinantes son iguales a 1. Las trazas (las sumas abajo de la diagonal principal) son igual a 2. Los valores característicos satisfacen $1 \cdot 1 = 1$ y $1 + 1 = 2$. Para T, B, y J, que son triangulares, los valores característicos están sobre la diagonal. Se quiere demostrar que *estas matrices son semejantes*: todas pertenecen a la misma familia.

(T) De T a J, la tarea es cambiar 2 a 1, y una matriz diagonal M lo hace:

$$M^{-1}TM = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = J.$$

(B) De B a J, la tarea es trasponer la matriz. Esto se obtiene con una permutación:

$$P^{-1}BP = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = J.$$

(A) De A a J, primero se va a T como en la ecuación (4). Luego se cambia 2 a 1:

$$U^{-1}AU = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = T \quad \text{y entonces} \quad M^{-1}TM = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = J.$$

Ejemplo 5 $A = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Cero es un triple valor característico de A y B, por lo que aparece en todos sus bloques de Jordan. Puede haber un solo bloque de 3 por 3, o un bloque de 2 por 2 y un bloque de 1 por 1, o tres bloques de 1 por 1. Así, A y B tienen tres posibles formas de Jordan:

$$J_1 = \begin{bmatrix} \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad J_2 = \begin{bmatrix} \mathbf{0} & \mathbf{1} & 0 \\ \mathbf{0} & \mathbf{0} & 0 \\ 0 & 0 & \mathbf{0} \end{bmatrix}, \quad J_3 = \begin{bmatrix} \mathbf{0} & 0 & 0 \\ 0 & \mathbf{0} & 0 \\ 0 & 0 & \mathbf{0} \end{bmatrix}. \tag{8}$$

El único vector característico de A es $(1, 0, 0)$. Su forma de Jordan sólo tiene un bloque, y A debe ser semejante a J_1. La matriz B tiene el vector característico adicional $(0, 1, 0)$, y su forma de Jordan es J_2 con dos bloques. Así como para J_3 = *matriz cero*, de suyo constituye una familia; la única matriz semejante a J_3 es $M^{-1}0M = 0$. Un conteo de los vectores característicos determina J cuando no hay nada más complicado que un valor característico triple.

Ejemplo 6 ***Aplicación a ecuaciones en diferencias y ecuaciones diferenciales*** *(potencias y exponenciales)*. Si A puede diagonalizarse, las potencias de $A = S\Lambda S^{-1}$ son fáciles: $A^k = S\Lambda^k S^{-1}$. En cada caso se tiene la semejanza de Jordan $A = MJM^{-1}$, de modo que ahora se requieren las potencias de J:

$$A^k = (MJM^{-1})(MJM^{-1})\cdots(MJM^{-1}) = MJ^kM^{-1}.$$

J es diagonal por bloques, y las potencias de cada uno de estos bloques pueden tomarse por separado:

$$(J_i)^k = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}^k = \begin{bmatrix} \lambda^k & k\lambda^{k-1} & \frac{1}{2}k(k-1)\lambda^{k-2} \\ 0 & \lambda^k & k\lambda^{k-1} \\ 0 & 0 & \lambda^k \end{bmatrix}. \tag{9}$$

Este bloque J_i entra en juego cuando λ es un triple valor característico con un solo vector característico. Su exponencial está en la solución de la ecuación diferencial correspondiente:

$$\textbf{Exponencial} \qquad e^{J_i t} = \begin{bmatrix} e^{\lambda t} & te^{\lambda t} & \frac{1}{2}t^2e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} \\ 0 & 0 & e^{\lambda t} \end{bmatrix}. \tag{10}$$

Aquí $I + J_it + (J_it)^2/2! + \cdots$ produce $1 + \lambda t + \lambda^2t^2/2! + \cdots = e^{\lambda t}$ en la diagonal.

La tercera columna de esta exponencial proviene directamente de resolver $du/dt = J_iu$:

$$\frac{d}{dt}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \quad \text{empezando con} \quad u_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Lo anterior puede resolverse por sustitución hacia atrás (ya que J_i es triangular). La última ecuación $du_3/dt = \lambda u_3$ produce $u_3 = e^{\lambda t}$. La ecuación para u_2 es $du_2/dt = \lambda u_2 + u_3$, y su solución es $te^{\lambda t}$. La ecuación de arriba es $du_1/dt = \lambda u_1 + u_2$, y su solución es $\frac{1}{2}t^2e^{\lambda t}$. Cuando λ tiene multiplicidad m con un solo vector característico, el factor adicional t aparece $m - 1$ veces.

Estas potencias y exponenciales de J forman parte de las soluciones u_k y $u(t)$. La otra parte es la M que relaciona la A original con la matriz más conveniente J:

$$\text{si} \quad u_{k+1} = Au_k \quad \text{entonces} \quad u_k = A^ku_0 = MJ^kM^{-1}u_0$$
$$\text{si} \quad du/dt = Au \quad \text{entonces} \quad u(t) = e^{At}u(0) = Me^{Jt}M^{-1}u(0).$$

Cuando M y J son S y Λ (el caso diagonalizable) aquéllas son las fórmulas de las secciones 5.3 y 5.4. El apéndice B regresa al caso no diagonalizable, y muestra cómo es posible alcanzar la forma de Jordan. Espero que la tabla siguiente sea un resumen conveniente.

Transformaciones de semejanza

1. A es ***diagonalizable***: las columnas de S son vectores característicos y $S^{-1}AS = \Lambda$.
2. A es ***arbitraria***: las columnas de M incluyen "vectores característicos generalizados" de A, y la forma de Jordan $M^{-1}AM = J$ es *diagonal por bloques*.
3. A es ***arbitraria***: la matriz unitaria U puede escogerse de tal manera que $U^{-1}AU = T$ es *triangular*.
4. A es ***normal***, $AA^H = A^HA$: luego, U puede escogerse de modo que $U^{-1}AU = \Lambda$.

Casos especiales de matrices normales, todas con vectores característicos ortonormales:

a) Si $A = A^H$ es hermitiana, entonces todos los λs son reales.
b) Si $A = A^T$ es simétrica real, entonces Λ es real y $U = Q$ es ortogonal.
c) Si $A = -A^H$ es antihermitiana, entonces todos los λ_i son puramente imaginarios.
d) Si A es ortogonal o unitaria, entonces todos los $|\lambda_i| = 1$ están en la circunferencia unitaria.

Conjunto de problemas 5.6

1. Si B es semejante a A y C es semejante a B, demuestre que C es semejante a A. (Sean $B = M^{-1}AM$ y $C = N^{-1}BN$.) ¿Qué matrices son semejantes a I?

2. Describa con palabras todas las matrices que son semejantes a $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, y encuentre dos de ellas.

3. Explique por qué A nunca es semejante a $A + I$.

4. Encuentre una diagonal M integrada por 1s y -1s, para demostrar que

$$A = \begin{bmatrix} 2 & 1 & & \\ 1 & 2 & 1 & \\ & 1 & 2 & 1 \\ & & 1 & 2 \end{bmatrix} \quad \text{es semejante a} \quad B = \begin{bmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 2 \end{bmatrix}.$$

5. Demuestre (si B es invertible) que BA es semejante a AB.

6. *a*) Si $CD = -DC$ (y D es invertible), demuestre que C es semejante a $-C$.
   b) Deduzca que los valores característicos de C deben presentarse por parejas más-menos.
   c) Demuestre directamente que si $Cx = \lambda x$, entonces $C(Dx) = -\lambda(Dx)$.

7. Considere cualquier A y una "rotación dada" M en el plano 1-2:

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}, \quad M = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Escoja el ángulo de rotación θ para obtener cero en el elemento (3, 1) de $M^{-1}AM$.

Nota Esta obtención "de ceros" no es fácil de continuar, ya que las rotaciones producen cero en lugar de d y h arruina el nuevo cero en la esquina. Es necesario dejar una diagonal abajo de la principal, y terminar el cálculo de los valores característicos de alguna otra forma. En caso contrario, si A puede hacerse diagonal y pueden verse sus valores característicos, entonces se encontrarían las raíces del polinomio det $(A - \lambda I)$ usando sólo las raíces cuadradas que determinan cos θ; lo cual es imposible.

8. ¿Qué matriz M cambia de la base $V_1 = (1, 1)$, $V_2 = (1, 4)$ a la base $v_1 = (2, 5)$, $v_2 = (1, 4)$? Las columnas de M se obtienen al expresar V_1 y V_2 como combinaciones $\Sigma m_{ij} v_i$ de los v's.

9. Para las dos mismas bases, exprese el vector (3, 9) como una combinación $c_1V_1 + c_2V_2$ y también como $d_1v_1 + d_2v_2$. Compruebe numéricamente que M relaciona c con d: $Mc = d$.

10. Confirme el último ejercicio: Si $V_1 = m_{11}v_1 + m_{21}v_2$ y $V_2 = m_{12}v_1 + m_{22}v_2$, y $m_{11}c_1 + m_{12}c_2 = d_1$ y $m_{21}c_1 + m_{22}c_2 = d_2$, los vectores $c_1V_1 + c_2V_2$ y $d_1v_1 + d_2v_2$ son los mismos. Esta es la "fórmula de cambio de base" $Mc = d$.

11. Si la transformación T es una reflexión, a través de la recta a 45° en el plano, encuentre su matriz respecto a la base estándar $v_1 = (1, 0)$, $v_2 = (0, 1)$, y también respecto a $V_1 = (1, 1)$, $V_2 = (1, -1)$. Demuestre que estas matrices son semejantes.

12. La *transformación identidad* lleva cada vector en sí mismo: $Tx = x$. Encuentre la matriz correspondiente, si la primera base es $v_1 = (1, 2)$, $v_2 = (3, 4)$ y la segunda base es $w_1 = (1, 0)$, $w_2 = (0, 1)$. (¡No es la matriz identidad!)

13. La derivada de $a + bx + cx^2$ es $b + 2cx + 0x^2$.
   a) Escriba la matriz de D3 por 3 tal que

$$D \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} b \\ 2c \\ 0 \end{bmatrix}.$$

   b) Calcule D^3, e interprete los resultados en términos de derivadas.
   c) ¿Cuáles son los valores característicos y los vectores característicos de D?

14. Demuestre que todo número es un vector característico para $Tf(x) = df/dx$, pero que la transformación $Tf(x) = \int_0^x f(t)\,dt$ no tiene valores característicos (aquí $-\infty < x < \infty$).

15. En el espacio de matrices de 2 por 2, sea T la transformación que *traspone cada matriz*. Encuentre los valores característicos y las "matrices características" para $A^{\mathrm{T}} = \lambda A$.

16. *a*) Encuentre una matriz ortogonal Q, de modo que $Q^{-1}AQ = \Lambda$ si

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{y} \quad \Lambda = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

   Luego, encuentre un segundo par de vectores ortonormales x_1, x_2 para $\lambda = 0$.
   b) Compruebe que $P = x_1x_1^{\mathrm{T}} + x_2x_2^{\mathrm{T}}$ es el mismo para ambos pares.

17. En dos pasos, demuestre que toda matriz *unitaria* A es diagonalizable:
   i) Si A es unitaria, y U también lo es, entonces $T = U^{-1}AU$ también es unitaria.
   ii) Una T triangular superior que es unitaria debe ser diagonal. Así, $T = \Lambda$.

   Cualquier matriz unitaria A (con valores característicos distintos o no) tiene un conjunto completo de vectores característicos ortonormales. Todos los valores característicos satisfacen $|\lambda| = 1$.

18. Encuentre una matriz normal ($NN^{\mathrm{H}} = N^{\mathrm{H}}N$) que no sea hermitiana, antihermitiana, unitaria, o diagonal. Demuestre que todas las matrices permutación son normales.

19. Suponga que T es una matriz triangular superior de 3 por 3, con elementos t_{ij}. Compare los elementos de TT^{H} y $T^{\mathrm{H}}T$, y demuestre que si son iguales, entonces T debe ser diagonal. Todas las matrices triangulares normales son diagonales.

20. Si N es normal, demuestre que $\|Nx\| = \|N^{\mathrm{H}}x\|$ para todo vector x. Deduzca que el i-ésimo renglón de N tiene la misma longitud que la i-ésima columna. *Nota*: Si N también es triangular superior, esto de nuevo lleva a la conclusión de que debe ser diagonal.

21. Demuestre que una matriz con vectores característicos ortonormales debe ser normal, como se establece en **5T**: Si $U^{-1}NU = \Lambda$, o $N = U\Lambda U^{\mathrm{H}}$, entonces $NN^{\mathrm{H}} = N^{\mathrm{H}}N$.

22. Encuentre una U unitaria y una T triangular de modo que $U^{-1}AU = T$, para

$$A = \begin{bmatrix} 5 & -3 \\ 4 & -2 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

23. Si A tiene valores característicos 0, 1, 2, ¿cuáles son los valores característicos de $A(A - I)(A - 2I)$?

24. *a*) Demuestre por multiplicación directa que toda matriz triangular T, por ejemplo de 3 por 3, satisface su propia ecuación característica $(T - \lambda_1 I)(T - \lambda_2 I)(T - \lambda_3 I) = 0$.
b) Sustituya $U^{-1}AU$ por T para deducir el famoso ***teorema de Cayley-Hamilton: Toda matriz satisface su propia ecuación característica***. Para las matrices de 3 por 3, esto es $(A - \lambda_1 I)(A - \lambda_2 I)(A - \lambda_3 I) = 0$.

25. El polinomio característico de $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ es $\lambda^2 - (a+d)\lambda + (ad - bc)$. Por sustitución directa, compruebe el teorema de Cayley-Hamilton: $A^2 - (a+d)A + (ad - bc)I = 0$.

26. Si $a_{ij} = 1$ arriba de la diagonal principal y $a_{ij} = 0$ en todas partes, encuentre la forma de Jordan (por ejemplo, para matrices de 4 por 4) encontrando todos los vectores característicos.

27. Demuestre al tanteo para una M que ningunas tomadas dos a dos de las tres formas de Jordan en la ecuación (8) son semejantes: $J_1 \neq M^{-1}J_2M$, $J_1 \neq M^{-1}J_3M$, y $J_2 \neq M^{-1}J_3M$.

28. Resuelva $u' = Ju$ por sustitución hacia atrás, resolviendo primero para $u_2(t)$:

$$\frac{du}{dt} = Ju = \begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad \text{con valor inicial} \quad u(0) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Observe te^{5t} en la primera componente $u_1(t)$.

29. Calcule A^{10} y e^A si $A = MJM^{-1}$:

$$A = \begin{bmatrix} 14 & 9 \\ -16 & -10 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -4 & 3 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 4 & 3 \end{bmatrix}.$$

30. Demuestre que A y B son semejantes encontrando una M tal que $B = M^{-1}AM$:

a) $A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$.

b) $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$.

c) $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ y $B = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$.

31. ¿Cuáles de las siguientes matrices A_1 a A_6 son semejantes?

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

32. Hay 16 matrices de 2 por 2 cuyos elementos son 0s y 1s. Las matrices semejantes pertenecen a la misma familia. ¿Cuántas familias hay? ¿Cuántas matrices (en total 16) hay en cada familia?

33. *a*) Si x está en el espacio nulo de A, demuestre que $M^{-1}x$ está en el espacio nulo de $M^{-1}AM$.
b) Los espacios nulos de A y $M^{-1}AM$ tienen igual(es)(vectores)(bases)(dimensión).

34. Si A y B tienen exactamente los mismos valores característicos y vectores característicos, ¿es $A = B$? Con n vectores característicos independientes, se tiene $A = B$. Encuentre $A \neq B$ cuando $\lambda = 0, 0$ (repetido), aunque sólo hay una línea de vectores característicos $(x_1, 0)$.

Los problemas 35 a 39 son sobre la forma de Jordan.

35. Por multiplicación directa, encuentre J^2 y J^3 cuando

$$J = \begin{bmatrix} c & 1 \\ 0 & c \end{bmatrix}.$$

Conjeture la forma de J^k. Haga $k = 0$ para obtener J^0. Haga $k = -1$ para obtener J^{-1}.

36. Si J es la matriz de Jordan por bloques de 5 por 5 con $\lambda = 0$, encuentre J^2 y cuente sus vectores característicos. También encuentre su forma de Jordan (dos bloques).

37. En el texto se resolvió $du/dt = Ju$ para una matriz J de Jordan por bloques de 3 por 3. Agregue una cuarta ecuación $dw/dt = 5w + x$. Siga el patrón de soluciones para z, y, x con la finalidad de encontrar w.

38. Los valores característicos de las siguientes matrices de Jordan son 0, 0, 0, 0. Las matrices tienen dos vectores característicos (encuéntrelos). Sin embargo, los tamaños de los bloques no coinciden y *J no es semejante a K*:

$$J = \left[\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] \quad \text{y} \quad K = \left[\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right].$$

Para cualquier matriz M, compare JM con MK. Si son iguales, demuestre que M no es invertible. Luego, $M^{-1}JM = K$ es imposible.

39. Demuestre en tres pasos que A^{T} *siempre es semejante a* A (se sabe que los λs son los mismos; el problema son los vectores característicos):
a) Para $A =$ un bloque, encuentre $M_i =$ permutación tal que $M_i^{-1}J_iM_i = J_i^{\mathrm{T}}$.
b) Para $A = J$ cualquiera, construya M_0, a partir de bloques, de modo que $M_0^{-1}JM_0 = J^{\mathrm{T}}$.
c) Para cualquiera $A = MJM^{-1}$, demuestre que A^{T} es semejante a J^{T} y que también es semejante a J y a A.

40. ¿Cuáles de los siguientes pares son semejantes? Escoja a, b, c, d para demostrar que los otros pares no lo son:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \begin{bmatrix} b & a \\ d & c \end{bmatrix} \quad \begin{bmatrix} c & d \\ a & b \end{bmatrix} \quad \begin{bmatrix} d & c \\ b & a \end{bmatrix}.$$

41. ¿Falso o verdadero (proporciona una buena razón)?
a) Una matriz invertible no puede ser semejante a una matriz singular.
b) Una matriz simétrica no puede ser semejante a una matriz no simétrica.
c) A no puede ser semejante a $-A$, a menos que $A = 0$.
d) $A - I$ no puede ser semejante a $A + I$.

42. Demuestre que *AB tiene los mismos valores característicos que BA*.

43. Si A es de 6 por 4 y B es de 4 por 6, AB y BA son de tamaños distintos. No obstante,

$$\begin{bmatrix} I & -A \\ 0 & I \end{bmatrix}\begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix}\begin{bmatrix} I & A \\ 0 & I \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix} = G.$$

a) ¿De qué tamaño son los bloques de G? Son los mismos en cada matriz.
b) Esta ecuación es $M^{-1}FM = G$, de modo que F y G tienen los mismos 10 valores característicos. F tiene los valores característicos de AB más 4 ceros; G tiene los

valores característicos de BA más 6 ceros. ***AB* tiene los mismos valores característicos que *BA* más _____ ceros.**

44. ¿Por qué cada una de las siguientes afirmaciones es verdadera?
a) Si A es semejante a B, entonces A^2 es semejante a B^2.
b) A^2 y B^2 pueden ser semejantes cuando A y B no son semejantes (intente $\lambda = 0, 0$).
c) $\begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix}$ es semejante a $\begin{bmatrix} 3 & 1 \\ 0 & 4 \end{bmatrix}$.
d) $\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$ no es semejante a $\begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$.
e) Si se intercambian los renglones 1 y 2 de A, y luego se intercambian las columnas 1 y 2, **los valores característicos siguen siendo los mismos.**

Propiedades de los valores característicos y vectores característicos

¿Cómo se reflejan las propiedades de una matriz en sus valores característicos y vectores característicos? Esta cuestión es fundamental en todo el capítulo 5. Una tabla que organice los hechos más importantes puede ser de utilidad. Para cada clase de matrices, a continuación se presentan las propiedades especiales de los valores característicos λ_i y los vectores característicos x_i.

Simétrica: $A^T = A$	λ real	ortogonal $x_i^T x_j = 0$
Ortogonal: $Q^T = Q^{-1}$	toda $\lvert\lambda\rvert = 1$	ortogonal $\bar{x}_i^T x_j = 0$
Simétrica sesgada: $A^T = -A$	λ imaginaria	ortogonal $\bar{x}_i^T x_j = 0$
Hermitiana compleja: $\bar{A}^T = A$	λ real	ortogonal $\bar{x}_i^T x_j = 0$
Definida positiva: $x^T A x > 0$	toda $\lambda > 0$	
Matriz similar: $B = M^{-1}AM$	$\lambda(B) = \lambda(A)$	$x(B) = M^{-1}x(A)$
Proyección: $P = P^2 = P^T$	$\lambda = 1;\ 0$	espacio de columna; espacio nulo
Reflexión: $I - 2uu^T$	$\lambda = -1;\ 1, \ldots, 1$	$u; u^{\perp}$
Matriz rango 1: uv^T	$\lambda = v^T u;\ 0, \ldots, 0$	$u;\ v^{\perp}$
Inversa: A^{-1}	$1/\lambda(A)$	vectores característicos de A
Corrimiento: $A + cI$	$\lambda(A) + c$	vectores característicos de A
Potencias estables: $A^n \to 0$	toda $\lvert\lambda\rvert < 1$	
Exponencialmente estable: $e^{At} \to 0$	toda Re $\lambda < 0$	
Markov: $m_{ij} > 0$, $\sum_{i=1}^{n} m_{ij} = 1$	$\lambda_{\text{máx}} = 1$	estado estacionario $x > 0$
Permutación cíclica: $P^n = I$	$\lambda_k = e^{2\pi ik/n}$	$x_k = (1, \lambda_k, \ldots, \lambda_k^{n-1})$
Diagonalizable: $S\Lambda S^{-1}$	diagonal de Λ	las columnas de S son independientes
Simétrica: $Q\Lambda Q^T$	diagonal de Λ (real)	las columnas de Q son ortonormales
Jordan: $J = M^{-1}AM$	diagonal de J	cada bloque de un vector característico
Cada matriz: $A = U\Sigma V^T$	rango (A) = rango (Σ)	los vectores característicos de A^TA, AA^T en V, U

Capítulo 5 Ejercicios de repaso

5.1 Encuentre los valores característicos, los vectores característicos, y la matriz de diagonalización S, para

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 7 & 2 \\ -15 & -4 \end{bmatrix}.$$

5.2 Encuentre los determinantes de A y A^{-1} si

$$A = S \begin{bmatrix} \lambda_1 & 2 \\ 0 & \lambda_2 \end{bmatrix} S^{-1}.$$

5.3 Si A tiene los valores característicos 0 y 1, correspondientes a los vectores característicos

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 2 \\ -1 \end{bmatrix},$$

¿podría decir de antemano que A es simétrica? ¿Cuáles son su traza y su determinante? ¿Cuál es A?

5.4 En el problema previo, ¿cuáles son los valores característicos y los vectores característicos de A^2? ¿Cuál es la relación de A^2 con A?

5.5 ¿Existe una matriz A tal que toda la familia $A + cI$ es invertible para todos los números complejos c? Encuentre una matriz real con $A + rI$ invertible para todo r real.

5.6 Resuelva para los dos valores iniciales, y luego encuentre e^{At}:

$$\frac{du}{dt} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} u \quad \text{si} \quad u(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{y si} \quad u(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

5.7 ¿Prefiere un interés compuesto trimestralmente a 40% anual o anualmente a 50%?

5.8 ¿Falso o verdadero? (Proporcione un contraejemplo si es falso):
a) Si B se forma a partir de A, mediante el intercambio de dos renglones, entonces B es semejante a A.
b) Si una matriz triangular es semejante a una matriz diagonal, ya es diagonal.
c) Cualesquiera de las dos afirmaciones anteriores implica la tercera: A es hermitiana, A es unitaria, $A^2 = I$.
d) Si A y B son diagonalizables, entonces también lo es AB.

5.9 ¿Qué ocurre a la sucesión de Fibonacci si se retrocede en el tiempo, y cómo está relacionado F_{-k} con F_k? La ley $F_{k+2} = F_{k+1} + F_k$ sigue siendo válida, de modo que $F_{-1} = 1$.

5.10 Encuentre la solución general de $du/dt = Au$ si

$$A = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

¿Puede encontrar un instante T al que se garantice que la solución $u(T)$ vuelva al valor inicial $u(0)$?

5.11 Si P es la matriz que proyecta $\mathbf{R}^n$ sobre un subespacio $\mathbf{S}$, explique por qué todo vector en $\mathbf{S}$ es un vector característico, y así como cada vector en $\mathbf{S}^\perp$. ¿Cuáles son los valores característicos? (Observe la relación con $P^2 = P$, lo que significa que $\lambda^2 = \lambda$.)

5.12 Demuestre que toda matriz de orden > 1 es la suma de dos matrices singulares.

5.13 *a*) Demuestre que la ecuación diferencial matricial $dX/dt = AX + XB$ tiene la solución $X(t) = e^{At}X(0)e^{Bt}$.
b) Demuestre que las soluciones de $dX/dt = AX - XA$ preservan los mismos valores característicos para todos los instantes.

5.14 Si los valores característicos de A son 1 y 3 con vectores característicos (5, 2) y (2, 1), encuentre las soluciones de $du/dt = Au$ y $u_{k+1} = Au_k$, empezando con $u = (9, 4)$.

5.15 Encuentre los valores característicos y los vectores característicos de

$$A = \begin{bmatrix} 0 & -i & 0 \\ i & 1 & i \\ 0 & -i & 0 \end{bmatrix}.$$

¿Qué propiedad se espera tengan los vectores característicos y es verdadera?

5.16 Intente resolver lo siguiente para demostrar que A no tiene raíz cuadrada.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = A$$

Cambie los elementos en la diagonal de A a 4 y encuentre una raíz cuadrada.

5.17 *a*) Encuentre los valores característicos y los vectores característicos de $A = \begin{bmatrix} 0 & 4 \\ \frac{1}{4} & 0 \end{bmatrix}$.
b) Resuelva $du/dt = Au$ empezando con $u(0) = (100, 100)$.
c) Si $v(t)$ = ganancia de los corredores de bolsa y $w(t)$ = ganancia del cliente, y ambos se ayudan mutuamente mediante $dv/dt = 4w$ y $dw/dt = \frac{1}{4}v$, ¿a que tiende la razón v/w cuando $t \to \infty$?

5.18 ¿Falso o verdadero? (Proporcione una razón si es verdadero y un contraejemplo si es falso).
a) Para toda matriz A, hay una solución de $du/dt = Au$ empezando con $u(0) = (1, \ldots, 1)$.
b) Toda matriz invertible puede diagonalizarse.
c) Toda matriz diagonalizable puede invertirse.
d) El intercambio de los renglones de una matriz de 2 por 2 invierte el signo de sus valores característicos.
e) Si los vectores característicos x y y corresponden a valores característicos distintos, entonces $x^{\mathrm{H}}y = 0$.

5.19 Si K es una matriz simétrica sesgada, demuestre que $Q = (I - K)(I + K)^{-1}$ es una matriz ortogonal. Encuentre Q si $K = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$.

5.20 Si $K^{\mathrm{H}} = -K$ (hermitiana sesgada), los valores característicos son imaginarios y los vectores característicos son ortogonales.
a) ¿Cómo se sabe que $K - I$ es invertible?
b) ¿Cómo se sabe que $K = U\Lambda U^{\mathrm{H}}$ para una matriz U unitaria?

c) ¿Por qué es unitaria $e^{\Lambda t}$?
d) ¿Por qué es unitaria e^{Kt}?

5.21 Si M es la matriz diagonal con elementos d, d^2, d^3, ¿cuál es $M^{-1}AM$? ¿Cuáles son sus valores característicos en el siguiente caso?

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

5.22 Si $A^2 = -I$, ¿cuáles son los valores característicos de A? Si A es una matriz real de n por n, demuestre que n debe ser par, y proporcione un ejemplo.

5.23 Si $Ax = \lambda_1 x$ y $A^{\mathrm{T}}y = \lambda_2 y$ (todos reales), demuestre que $x^{\mathrm{T}}y = 0$.

5.24 Una variante de la matriz de Fourier es la "matriz seno":

$$S = \frac{1}{\sqrt{2}}\begin{bmatrix} \operatorname{sen}\theta & \operatorname{sen}2\theta & \operatorname{sen}3\theta \\ \operatorname{sen}2\theta & \operatorname{sen}4\theta & \operatorname{sen}6\theta \\ \operatorname{sen}3\theta & \operatorname{sen}6\theta & \operatorname{sen}9\theta \end{bmatrix} \quad \text{con} \quad \theta = \frac{\pi}{4}.$$

Compruebe que $S^{\mathrm{T}} = S^{-1}$. (Las columnas son los vectores característicos de la matriz tridiagonal $-1, 2, -1$).

5.25 *a*) Encuentre una matriz N diferente de cero tal que $N^3 = 0$.
b) Si $Nx = \lambda x$, demuestre que λ debe ser cero.
c) Demuestre que N (denominada "matriz nilpotente") no puede ser simétrica.

5.26 a) Encuentre la matriz $P = aa^{\mathrm{T}}/a^{\mathrm{T}}a$ que proyecta cualquier vector sobre la recta que pasa por $a = (2, 1, 2)$.
b) ¿Cuál es el único valor característico diferente de cero de P, y cuál es el vector característico correspondiente?
c) Resuelva $u_{k+1} = Pu_k$, empezando con $u_0 = (9, 9, 0)$.

5.27 Suponga que el primer renglón de A es 7, 6 y que sus valores característicos son $i, -i$. Encuentre A.

5.28 *a*) ¿Para qué números c y d ocurre que A tiene valores característicos reales y vectores característicos ortogonales?

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & d & c \\ 0 & 5 & 3 \end{bmatrix}.$$

b) ¿Para cuáles c y d es posible encontrar tres vectores ortonormales que sean combinaciones de las columnas? (¡No lo resuelva!)

5.29 Si los vectores x_1 y x_2 están en las columnas de S, ¿cuáles son los valores característicos y los vectores característicos de

$$A = S\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}S^{-1} \quad \text{y} \quad B = S\begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}S^{-1}?$$

5.30 ¿Cuál es el límite cuando $k \to \infty$ (el estado estacionario de Markov) de $\begin{bmatrix} 0.4 & 0.3 \\ 0.6 & 0.7 \end{bmatrix}^k \begin{bmatrix} a \\ b \end{bmatrix}$?

Capítulo



Matrices positivas definidas

6.1 MÍNIMOS, MÁXIMOS, Y PUNTOS SILLA

Hasta el momento, difícilmente se ha pensado en **los signos de los valores característicos**. No es posible preguntar si λ es positivo antes de saber si es real. En el capítulo 5 se estableció que toda matriz simétrica tiene valores característicos reales. A continuación se encontrará una prueba que puede aplicarse directamente a A, sin calcular sus valores característicos, ***que garantizará que todos estos valores característicos son positivos***. Esta prueba conlleva tres de los conceptos más importantes del libro: *pivotes*, *determinantes* y *valores característicos*.

A menudo, el signo de los valores característicos es crucial. Para estabilidad en ecuaciones diferenciales, se requieren valores característicos negativos, de modo que $e^{\lambda t}$ decaiga. El nuevo y muy importante nuevo problema es reconocer un ***punto mínimo***. Esto se presenta en toda la ciencia e ingeniería, así como en todo problema de optimización. El problema matemático es mover la prueba de la segunda derivada $F'' > 0$ hacia n dimensiones. A continuación se presentan dos ejemplos:

$$F(x, y) = 7 + 2(x + y)^2 - y \operatorname{sen} y - x^3 \qquad f(x, y) = 2x^2 + 4xy + y^2.$$

Ya sea $F(x, y)$ o $f(x, y)$, ¿tiene un punto mínimo en $x = y = 0$?

Observación 1 Los *términos de orden cero* $F(0, 0) = 7$ y $f(0, 0) = 0$ no afectan la respuesta. Simplemente suben o bajan las gráficas de F y f.

Observación 2 Los *términos lineales* proporcionan una condición necesaria: A fin de tener alguna posibilidad de un mínimo, las primeras derivadas deben hacerse cero en $x = y = 0$:

$$\frac{\partial F}{\partial x} = 4(x + y) - 3x^2 = 0 \quad \text{y} \quad \frac{\partial F}{\partial y} = 4(x + y) - y \cos y - \operatorname{sen} y = 0$$

$$\frac{\partial f}{\partial x} = 4x + 4y = 0 \quad \text{y} \quad \frac{\partial f}{\partial y} = 4x + 2y = 0. \quad \textit{Todo cero.}$$

Así, $(x, y) = (0, 0)$ es un *punto estacionario* para ambas funciones. La superficie $z = F(x, y)$ es tangente al plano horizontal $z = 7$, y la superficie $z = f(x, y)$ es tangente al plano $z = 0$. La cuestión es si las gráficas *pasan o no por arriba de esos planos*, a medida que se aleja el punto de tangencia $x = y = 0$.

Observación 3 ***Las segundas derivadas en* (0, 0) *son decisivas:***

$$\frac{\partial^2 F}{\partial x^2} = 4 - 6x = 4 \qquad \frac{\partial^2 f}{\partial x^2} = 4$$

$$\frac{\partial^2 F}{\partial x \partial y} = \frac{\partial^2 F}{\partial y \partial x} = 4 \qquad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 4$$

$$\frac{\partial^2 F}{\partial y^2} = 4 + y \operatorname{sen} y - 2\cos y = 2 \qquad \frac{\partial^2 f}{\partial y^2} = 2.$$

Estas segundas derivadas 4, 4, 2 contienen la respuesta. Debido a que son las mismas para F y para f, deben contener la misma respuesta. Las dos funciones se comportan exactamente de la misma manera cerca del origen. ***F tiene un mínimo si y sólo si f tiene un mínimo.*** ¡Se demostrará que estas funciones no lo hacen!

Observación 4 Los *términos de orden superior* en F no afectan la cuestión de un mínimo *local*, aunque pueden impedir que éste se convierta en un mínimo *global*. En el ejemplo, el término $-x^3$ tarde o temprano empujará a F hacia $-\infty$. Para $f(x, y)$, sin términos superiores, toda la acción está en (0, 0).

Toda forma cuadrática $f = ax^2 + 2bxy + cy^2$ *tiene un punto estacionario en el origen, donde* $\partial f/\partial x = \partial f/\partial y = 0$. Un mínimo local también debe ser un mínimo global. Entonces, la forma de la superficie $z = f(x, y)$ es como la de un tazón, apoyado en el origen (véase la figura 6.1). Si el punto estacionario de F está en $x = \alpha$, $y = \beta$, el único cambio es en el uso de las segundas derivadas en α, β:

Parte cuadrática de F

$$f(x, y) = \frac{x^2}{2}\frac{\partial^2 F}{\partial x^2}(\alpha, \beta) + xy\frac{\partial^2 F}{\partial x \partial y}(\alpha, \beta) + \frac{y^2}{2}\frac{\partial^2 F}{\partial y^2}(\alpha, \beta). \qquad (1)$$

Cerca de (0, 0), está $f(x, y)$, y se comporta de la misma forma en que $F(x, y)$ se comporta cerca de (α, β).

Las terceras derivadas hacen su aparición en el problema cuando las segundas derivadas fracasan en proporcionar una decisión definitiva. Esto ocurre cuando la parte cuadrática es singular. Para un mínimo verdadero, se permite que f sea cero *sólo en* $x = y = 0$.



Figura 6.1 Tazón y silla: $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ definida y $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ indefinida.

Cuando $f(x, y)$ es estrictamente positiva en todos los demás puntos (el cuenco del tazón está hacia arriba), se denomina ***positiva definida.***

Definida contra indefinida: tazón contra silla

El problema se reduce a lo siguiente: Para una función de dos variables x y y, ¿cuál es la sustitución correcta para la condición $\partial^2 F/\partial x^2 > 0$? Con una sola variable, el signo de la segunda derivada decide entre un mínimo o un máximo. Ahora se cuenta con tres segundas derivadas: F_{xx}, $F_{xy} = F_{yx}$, y F_{yy}. Estos tres números (como 4, 4, 2) deben determinar si F (igual que f) tiene o no un mínimo.

¿Qué condiciones sobre a, b y c aseguran que la cuadrática $f(x, y) = ax^2 + 2bxy + cy^2$ es positiva definida? Una condición necesaria es fácil:

i) *Si* $ax^2 + 2bxy + cy^2$ *es positiva definida, entonces necesariamente* $\boldsymbol{a > 0}$.

Se considera $x = 1$, $y = 0$, donde $ax^2 + 2bxy + cy^2$ es igual a a. Ésta debe ser positiva. Trasladando de vuelta a F, lo anterior significa que $\partial^2 F/\partial x^2 > 0$. La gráfica debe ser hacia arriba en la dirección x. De manera semejante, se fija $x = 0$ y se considera la dirección y, donde $f(0, y) = cy^2$:

ii) *Si* $f(x, y)$ *es positiva definida, entonces necesariamente* $\boldsymbol{c > 0}$.

¿Estas condiciones sobre $a > 0$ y $c > 0$ garantizan que $f(x, y)$ siempre es positiva? La respuesta es **no**. Un gran término cruzado $2bxy$ puede empujar la gráfica por abajo de cero.

Ejemplo 1 $f(x, y) = x^2 - 10xy + y^2$. Aquí $a = 1$ y $c = 1$ son ambas positivas. Sin embargo, f no es positiva definida, ya que $f(1, 1) = -8$. Las condiciones $a > 0$ y $c > 0$ aseguran que $f(x, y)$ es positiva en los ejes x y y. Pero la función es negativa sobre la recta $x = y$, porque $b = -10$ supera a a y c.

Ejemplo 2 En la f original, el coeficiente $2b = 4$ era positivo. ¿Esto asegura un mínimo? De nuevo la respuesta es **no**; ¡el signo de b carece de importancia! *Aunque sus segundas derivadas son positivas,* $2x^2 + 4xy + y^2$ *no es positiva definida.* ***Ni F ni f tienen un mínimo en* (0, 0) *porque* $\boldsymbol{f(1, -1) = 2 - 4 + 1 = -1}$.**

Lo que debe controlarse es el tamaño de b, en comparación con a y c. Ahora se quiere una condición necesaria y suficiente para la condición de positiva definida. La técnica más simple es completar el cuadrado:

Expresar $f(x, y)$ usando cuadrados

$$f = ax^2 + 2bxy + cy^2 = a\left(x + \frac{b}{a}y\right)^2 + \left(c - \frac{b^2}{a}\right)y^2. \qquad (2)$$

El primer término a la derecha nunca es negativo, cuando el cuadrado se multiplica por $a > 0$. Pero este cuadrado puede ser cero, y entonces el segundo término debe ser positivo. El coeficiente de ese término es $(ac - b^2)/a$. El último requerimiento para la condición de positiva definida es que este coeficiente debe ser positivo:

iii) *Si* $ax^2 + 2bxy + cy^2$ *permanece positiva, entonces necesariamente* $\boldsymbol{ac > b^2}$.

Prueba para un mínimo: Las condiciones $a > 0$ y $ac > b^2$ son correctas. Garantizan $c > 0$. El miembro derecho de (2) es positivo, y se ha encontrado un mínimo:

6A $ax^2 + 2bxy + cy^2$ es positiva definida si y sólo si $a > 0$ y $ac > b^2$. Cualquier $F(x, y)$ tiene un mínimo en un punto donde $\partial F/\partial x = \partial F/\partial y = 0$ con

$$\frac{\partial^2 F}{\partial x^2} > 0 \quad \text{y} \quad \left[\frac{\partial^2 F}{\partial x^2}\right]\left[\frac{\partial^2 F}{\partial y^2}\right] > \left[\frac{\partial^2 F}{\partial x \partial y}\right]^2. \tag{3}$$

Prueba para un máximo: Debido a que f tiene un máximo siempre que $-f$ tiene un mínimo, simplemente se invierten los signos de a, b y c. Esto en realidad deja $ac > b^2$ sin cambio: la forma cuadrática es ***negativa definida*** si y sólo si $a < 0$ y $ac > b^2$. El mismo cambio es válido para un máximo de $F(x, y)$.

Caso singular $ac = b^2$: El segundo término en la ecuación (2) desaparece para dejar sólo el primer cuadrado, que es ***positivo semidefinido***, cuando $a > 0$, o ***negativo semidefinido***, cuando $a < 0$. El prefijo *semi* permite la posibilidad de que F pueda ser igual a cero, como es el caso en el punto $x = b$, $y = -a$. La superficie $z = f(x, y)$ degenera de un tazón en un valle. Para $f = (x + y)^2$, el valle se encuentra a lo largo de la recta $x + y = 0$.

Punto silla $ac < b^2$: En una dimensión, $F(x)$ tiene un mínimo o un máximo, o bien, $F'' = 0$. En dos dimensiones, permanece una posibilidad bastante importante: *la combinación $ac - b^2$ puede ser negativa.* Esto ocurrió en los dos ejemplos, cuando b dominaba a a y c. Esto también ocurre si a y c tienen signos opuestos. Así, dos direcciones proporcionan resultados opuestos: en una dirección, f crece; en la otra, decrece. Resulta de utilidad considerar dos casos especiales:

$$\textbf{Puntos silla en (0, 0)} \quad f_1 = 2xy \quad \text{y} \quad f_2 = x^2 - y^2 \quad \text{y} \quad ac - b^2 = -1.$$

En el primero, $b = 1$ domina $a = c = 0$. En el segundo, $a = 1$ y $c = -1$ tienen signos opuestos. Los puntos silla $2xy$ y $x^2 - y^2$ son prácticamente los mismos; si uno se hace girar 45°, se obtiene el otro. También es difícil trazarlos.

Estas formas cuadráticas son ***indefinidas***, porque pueden asumir cualquier signo. Así, se tiene un punto estacionario que no es máximo ni mínimo. Se denomina ***punto silla***. La superficie $z = x^2 - y^2$ va hacia abajo en la dirección del eje y, donde sus piernas se ajustan (si usted ha montado a caballo). En caso de que usted haya cambiado a un automóvil, piense en una carretera que va por un paso de montaña. La cima del paso es un mínimo mientras se observa a lo largo de la cordillera, aunque es un máximo cuando usted avanza por la carretera.

Dimensiones superiores: álgebra lineal

El cálculo podría bastar para encontrar nuestras condiciones $F_{xx} > 0$ y $F_{xx}F_{yy} > F_{xy}^2$ para un mínimo. Sin embargo, el álgebra lineal está preparada para hacer más, ya que las segundas derivadas se ajustan a una matriz simétrica A. Los términos ax^2 y cy^2 aparecen *sobre la diagonal*. La derivada cruzada $2bxy$ está separada entre el mismo elemento b arriba y abajo. ¡Una cuadrática $f(x, y)$ proviene directamente de una matriz simétrica de 2 por 2!

$$x^{\mathrm{T}}Ax \text{ en } \mathbf{R}^2 \qquad ax^2 + 2bxy + cy^2 = \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}. \tag{4}$$

Esta identidad (por favor, realice el producto) constituye la clave para todo el capítulo. Se generaliza de inmediato a n dimensiones, y constituye una abreviación perfecta para estudiar máximos y mínimos. Cuando las variables son $x_1, \ldots, x_n$, van en un vector columna x. ***Para cualquier matriz simétrica A, el producto $x^{\mathrm{T}}Ax$ es una forma cuadrática pura $f(x_1, \ldots, x_n)$:***

$$x^{\mathrm{T}}Ax \text{ en } \mathbf{R}^n \qquad \begin{bmatrix} x_1 & x_2 & \cdot & x_n \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdot & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ x_n \end{bmatrix} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j. \tag{5}$$

Los elementos diagonales a_{11} a a_{nn} multiplican x_1^2 y x_n^2. El par $a_{ij} = a_{ji}$ se combina en $2a_{ij}x_ix_j$. Así, $f = a_{11}x_1^2 + 2a_{12}x_1x_2 + \cdots + a_{nn}x_n^2$.

No hay términos de orden superior ni términos de orden inferior; sólo de segundo orden. La función es cero en $x = (0, \ldots, 0)$, y sus primeras derivadas son cero. La tangente es plana; se trata de un punto estacionario. Es necesario decidir si $x = 0$ es un mínimo, un máximo o un punto silla de la función $f = x^{\mathrm{T}}Ax$.

Ejemplo 3 $f = 2x^2 + 4xy + y^2$ y $A = \begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix} \rightarrow$ *punto silla.*

Ejemplo 4 $f = 2xy$ y $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \rightarrow$ *punto silla.*

Ejemplo 5 A es de 3 por 3 para $2x_1^2 - 2x_1x_2 + 2x_2^2 - 2x_2x_3 + 2x_3^2$:

$$f = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \rightarrow \textit{mínimo en } (0, 0, 0).$$

Cualquier función $F(x_1, \ldots, x_n)$ es aproximada de la misma forma. En un punto estacionario todas las primeras derivadas son cero. A es la **"matriz segunda derivada"** con elementos $a_{ij} = \partial^2 F/\partial x_i \partial x_j$. Esto automáticamente es igual a $a_{ji} = \partial^2 F/\partial x_j \partial x_i$, de modo que A es simétrica. **Así, *F tiene un mínimo cuando la cuadrática pura $x^{\mathrm{T}}Ax$ es positiva definida.*** Estos términos de segundo orden controlan a F cerca del punto estacionario:

$$\textbf{Serie de Taylor} \quad F(x) = F(0) + x^{\mathrm{T}}(\text{grad } F) + \frac{1}{2}x^{\mathrm{T}}Ax + \text{términos de orden superior.} \tag{6}$$

En un punto estacionario, $F = (\partial F/\partial x_1, \ldots, \partial F/\partial x_n)$ es un vector de ceros. Las segundas derivadas *en* $x^{\mathrm{T}}Ax$ asumen la gráfica hacia arriba, hacia abajo (o en una silla). Si el punto estacionario está en x_0 en vez de en 0, entonces $F(x)$ y todas las derivadas se calculan en x_0. Luego, x cambia a $x - x_0$ en el miembro derecho.

La siguiente sección contiene las pruebas para decidir si $x^{\mathrm{T}}Ax$ es positiva (el tazón se dirige hacia arriba a partir de $x = 0$). De manera equivalente, **las pruebas deciden si la matriz A es positiva definida**, lo cual constituye el objetivo más importante del capítulo.

Conjunto de problemas 6.1

1. La cuadrática $f = x^2 + 4xy + 2y^2$ tiene un punto silla en el origen, a pesar de que sus coeficientes son positivos. Escriba f como una *diferencia de dos cuadrados.*

2. Decida a favor o en contra de la condición de positiva definida de las siguientes matrices, y escriba las $f = x^TAx$ correspondientes:

$$a)\ \begin{bmatrix}1 & 3\\3 & 5\end{bmatrix}. \quad b)\ \begin{bmatrix}1 & -1\\-1 & 1\end{bmatrix}. \quad c)\ \begin{bmatrix}2 & 3\\3 & 5\end{bmatrix}. \quad d)\ \begin{bmatrix}-1 & 2\\2 & -8\end{bmatrix}.$$

El determinante en el inciso *b*) es cero; ¿a lo largo de cuál recta se encuentra $f(x, y) = 0$?

3. Si una matriz simétrica de 2 por 2 satisface las pruebas $a > 0$, $ac > b^2$, resuelva la ecuación cuadrática det $(A - \lambda I) = 0$, y demuestre que los dos valores característicos son positivos.

4. Decida entre un mínimo, un máximo o un punto silla para las siguientes funciones.
   a) $F = -1 + 4(e^x - x) - 5x \operatorname{sen} y + 6y^2$ en el punto $x = y = 0$.
   b) $F = (x^2 - 2x)\cos y$, con punto estacionario en $x = 1$, $y = \pi$.

5. *a*) ¿Para cuáles números b se cumple que la matriz $A = \begin{bmatrix}1 & b\\ b & 9\end{bmatrix}$ es positiva definida?
   b) Factorice $A = LDL^T$ cuando b está en el intervalo para la propiedad de positiva definida.
   c) Encuentre un valor mínimo de $\frac{1}{2}(x^2 + 2bxy + 9y^2) - y$ para b en este intervalo.
   d) ¿Cuál es el mínimo si $b = 3$?

6. Suponga que los coeficientes positivos a y c dominan a b en el sentido de que $a + c > 2b$. Encuentre un ejemplo que tenga $ac < b^2$, de modo que la matriz no sea positiva definida.

7. *a*) ¿Cuáles matrices simétricas de 3 por 3 A_1 y A_2 corresponden a f_1 y f_2?

$$f_1 = x_1^2 + x_2^2 + x_3^2 - 2x_1x_2 - 2x_1x_3 + 2x_2x_3$$

$$f_2 = x_1^2 + 2x_2^2 + 11x_3^2 - 2x_1x_2 - 2x_1x_3 - 4x_2x_3.$$

   b) Demuestre que f_1 es una matriz cuadrado perfecto *simple* y no positiva definida. ¿Dónde ocurre que f_1 es igual a 0?
   c) Factorice A_2 en LL^T. Escriba $f_2 = x^TA_2x$ como una suma de tres cuadrados.

8. Si $A = \begin{bmatrix}a & b\\ b & c\end{bmatrix}$ es positiva definida, pruebe $A^{-1} = \begin{bmatrix}p & q\\ q & r\end{bmatrix}$ en cuanto a esta propiedad.

9. La cuadrática $f(x_1, x_2) = 3(x_1 + 2x_2)^2 + 4x_2^2$ es positiva. Encuentre su matriz A, factorícela en LDL^T, y relacione los elementos en D y L con 3, 2, 4 en f.

10. Si $R = \begin{bmatrix}p & q\\ q & r\end{bmatrix}$, escriba R^2 y compruebe que es positiva definida a menos que R sea singular.

11. *a*) Si $A = \begin{bmatrix}a & b\\ \bar{b} & c\end{bmatrix}$ es hermitiana (b *complejo*), encuentre sus pivotes y su determinante.
   b) Complete el cuadrado para x^HAx. Ahora $x^H = \begin{bmatrix}\bar{x}_1 & \bar{x}_2\end{bmatrix}$ puede ser compleja.

$$a|x_1|^2 + 2\operatorname{Re} b\bar{x}_1x_2 + c|x_2|^2 = a|x_1 + (b/a)x_2|^2 + ______|x_2|^2.$$

   c) Demuestre que $a > 0$ y $ac > |b|^2$ aseguran que A es positiva definida.
   d) Las matrices $\begin{bmatrix}1 & 1+i\\ 1-i & 2\end{bmatrix}$ y $\begin{bmatrix}3 & 4+i\\ 4-i & 6\end{bmatrix}$, ¿son positivas definidas?

12. Decida si $F = x^2y^2 - 2x - 2y$ tiene un mínimo en el punto $x = y = 1$ (después de demostrar que las primeras derivadas son cero en ese punto).

13. ¿En qué condiciones sobre a, b, c se cumple que $ax^2 + 2bxy + cy^2 > x^2 + y^2$ para toda x, y?

Los problemas 14 a 18 son sobre pruebas para la propiedad de positiva definida.

14. ¿Cuáles de A_1, A_2, A_3, A_4 tienen dos valores característicos positivos? Pruebe $a > 0$ y $ac > b^2$; no calcule los valores característicos. Encuentre una x, de modo que $x^TA_1x < 0$.

$$A_1 = \begin{bmatrix}5 & 6\\6 & 7\end{bmatrix} \quad A_2 = \begin{bmatrix}-1 & -2\\-2 & -5\end{bmatrix} \quad A_3 = \begin{bmatrix}1 & 10\\10 & 100\end{bmatrix} \quad A_4 = \begin{bmatrix}1 & 10\\10 & 101\end{bmatrix}.$$

15. ¿Cuál es la cuadrática $f = ax^2 + 2bxy + cy^2$ para cada una de las siguientes matrices? Complete el cuadrado con la finalidad de escribir f como una suma de uno o dos cuadrados $d_1(\ \ )^2 + d_2(\ \ )^2$.

$$A = \begin{bmatrix}1 & 2\\2 & 9\end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix}1 & 3\\3 & 9\end{bmatrix}.$$

16. Demuestre que $f(x, y) = x^2 + 4xy + 3y^2$ no tiene un mínimo en (0, 0) incluso si sus coeficientes son positivos. Escriba f como una *diferencia* de cuadrados y encuentre un punto (x, y) donde f es negativa.

17. (*Importante*) Si A tiene *columnas independientes*, entonces A^TA es cuadrada, simétrica e invertible (véase la sección 4.2). ***Vuelva a escribir x^TA^TAx para demostrar por qué es positiva, excepto cuando $x = 0$.*** Entonces, A^TA es positiva definida.

18. Pruebe si A^TA es positiva definida en cada caso:

$$A = \begin{bmatrix}1 & 2\\0 & 3\end{bmatrix}, \quad A = \begin{bmatrix}1 & 1\\1 & 2\\2 & 1\end{bmatrix}, \quad \text{y} \quad A = \begin{bmatrix}1 & 1 & 2\\1 & 2 & 1\end{bmatrix}.$$

19. Encuentre la matriz A de 3 por 3, sus pivotes, rango, valores característicos, y determinante:

$$\begin{bmatrix}x_1 & x_2 & x_3\end{bmatrix}\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = 4(x_1 - x_2 + 2x_3)^2.$$

20. Para $F_1(x, y) = \frac{1}{4}x^4 + x^2y + y^2$ y $F_2(x, y) = x^3 + xy - x$, encuentre las matrices segunda derivada A_1 y A_2:

$$A = \begin{bmatrix}\partial^2F/\partial x^2 & \partial^2F/\partial x\partial y\\ \partial^2F/\partial y\partial x & \partial^2F/\partial y^2\end{bmatrix}.$$

A_1 es positiva definida, de modo que F_1 es cóncava hacia arriba (= convexa). Encuentre el punto mínimo de F_1 y el punto silla de F_2 (analice, dónde las primeras derivadas son cero).

21. La gráfica de $z = x^2 + y^2$ es un tazón que se abre hacia arriba. *La gráfica de $z = x^2 - y^2$ es una silla.* La gráfica de $z = -x^2 - y^2$ es un tazón que se abre hacia abajo. ¿Cuál es una prueba sobre $F(x, y)$ para tener una silla en (0, 0)?

22. ¿Con qué valores de c se obtiene un tazón, y con cuáles un punto silla para la gráfica de $z = 4x^2 + 12xy + cy^2$? Describa esta gráfica en el valor fronterizo de c.

6.2 PRUEBAS PARA COMPROBAR SI UNA MATRIZ ES POSITIVA DEFINIDA

¿Cuáles son las matrices simétricas que poseen la propiedad de que $x^TAx > 0$ para todos los vectores x diferentes de cero? Hay cuatro o cinco formas distintas para contestar esta pregunta, y esperamos encontrarlas todas. La sección previa empezó con algunas sugerencias sobre los signos de los valores característicos, aunque ello originó las pruebas sobre a, b, c:

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \quad \text{es positiva definida cuando } a > 0 \quad \text{y} \quad ac - b^2 > 0.$$

A partir de estas condiciones, ***ambos valores característicos son positivos.*** Su producto $\lambda_1\lambda_2$ es el determinante $ac - b^2 > 0$, de modo que los valores característicos son ambos positivos o ambos negativos. Deben ser positivos porque su suma es la traza $a + c > 0$.

Al considerar a y $ac - b^2$, incluso es posible pronosticar la apariencia de los ***pivotes***. Éstos aparecieron cuando x^TAx se descompuso en una suma de cuadrados:

Suma de cuadrados $$ax^2 + 2bxy + cy^2 = a\left(x + \frac{b}{a}y\right)^2 + \frac{ac - b^2}{a}y^2. \qquad (1)$$

Estos coeficientes a y $(ac - b^2)/a$ son los pivotes para una matriz de 2 por 2. Para matrices más grandes, los pivotes siguen constituyendo una prueba para comprobar si una matriz es positiva definida: x^TAx es positiva cuando n cuadrados independientes se multiplican por ***pivotes positivos***.

Una observación preliminar más. Las dos partes de este libro se vincularon mediante el capítulo sobre determinantes. En consecuencia, se pregunta cuál es el papel que desempeñan los determinantes. ***No basta requerir que el determinante de A sea positivo.*** Si $a = c = -1$ y $b = 0$, entonces $\det A = 1$, pero $A = -I$ = negativa definida. La prueba del determinante se aplica no sólo a A misma, para obtener $ac - b^2 > 0$, sino también a la submatriz a de 1 por 1 que está en la esquina superior izquierda.

La generalización natural implica a todas las n de las *submatrices superiores izquierdas* de A:

$$A_1 = [a_{11}], \quad A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \ldots, \quad A_n = A.$$

A continuación se presenta el teorema principal sobre la comprobación de si una matriz es positiva definida, así como una demostración razonablemente detallada:

6B Cada una de las pruebas siguientes constituye una prueba necesaria y suficiente para que la matriz simétrica real A sea ***positiva definida***:

(I) $x^TAx > 0$ para todos los vectores reales x diferentes de cero.
(II) Todos los valores característicos de A satisfacen $\lambda_i > 0$.
(III) Todas las submatrices superiores izquierdas A_k tienen determinantes positivos.
(IV) Todos los pivotes (sin intercambios de renglones) cumplen $d_k > 0$.

Demostración La condición I define una matriz positiva definida. El primer paso es demostrar que cada valor característico es positivo:

$$\text{Si } Ax = \lambda x, \text{ entonces } x^TAx = x^T\lambda x = \lambda\|x\|^2.$$

Una matriz positiva definida tiene valores característicos positivos, ya que $x^TAx > 0$.

A continuación se procede en la otra dirección. Si todos los $\lambda_i > 0$, es necesario demostrar que $x^TAx > 0$ para todo vector x (no sólo para los vectores característicos). Debido a que las matrices simétricas tienen un conjunto completo de vectores característicos ortonormales, cualquier x es una combinación $c_1x_1 + \cdots + c_nx_n$. Luego,

$$Ax = c_1Ax_1 + \cdots + c_nAx_n = c_1\lambda_1x_1 + \cdots + c_n\lambda_nx_n.$$

Debido a la ortogonalidad, $x_i^Tx_j = 0$, y la normalización $x_i^Tx_i = 1$,

$$x^TAx = \left(c_1x_1^T + \cdots + c_nx_n^T\right)(c_1\lambda_1x_1 + \cdots + c_n\lambda_nx_n)$$
$$= c_1^2\lambda_1 + \cdots + c_n^2\lambda_n. \qquad (2)$$

Si todo $\lambda_i > 0$, entonces la ecuación (2) demuestra que $x^TAx > 0$. Así, la condición II implica la condición I.

Si la condición I se cumple, entonces también se cumple la condición III: El determinante de A es el producto de los valores característicos. Y si la condición I se cumple, ya se sabe que estos valores característicos son positivos. Pero también debe tratarse con toda submatriz superior izquierda A_k. El truco consiste en considerar a todos los vectores diferentes de cero cuyas últimas $n - k$ componentes sean cero:

$$x^TAx = \begin{bmatrix} x_k^T & 0 \end{bmatrix} \begin{bmatrix} A_k & * \\ * & * \end{bmatrix} \begin{bmatrix} x_k \\ 0 \end{bmatrix} = x_k^TA_kx_k > 0.$$

Así, A_k es positiva definida. Sus valores característicos (¡no los mismos λ_i!) deben ser positivos. Su determinante es su producto, por lo que todos los determinantes izquierdos superiores son positivos.

Si la condición III se cumple, entonces también se cumple la condición IV: Según la sección 4.4, el k-ésimo pivote d_k es la razón de $\det A_k$ a $\det A_{k-1}$. Si todos los determinantes son positivos, también lo son los pivotes.

Si la condición IV se cumple, entonces también se cumple la condición I: Se cuenta con pivotes positivos, y es necesario deducir que $x^TAx > 0$. Esto fue lo que se hizo en el caso de 2 por 2 al completar el cuadrado. Los pivotes eran los números fuera de los cuadrados. Para ver cómo ocurre lo anterior para matrices simétricas de cualquier tamaño, se regresa a la *eliminación de una matriz simétrica*: $A = LDL^T$.

Ejemplo 1 Pivotes positivos 2, $\frac{3}{2}$ y $\frac{4}{3}$:

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{bmatrix} \begin{bmatrix} 2 & & \\ & \frac{3}{2} & \\ & & \frac{4}{3} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 1 \end{bmatrix} = LDL^T.$$

Se busca separar x^TAx en x^TLDL^Tx:

$$\text{Si } x = \begin{bmatrix} u \\ v \\ w \end{bmatrix}, \text{ entonces } L^Tx = \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} u - \frac{1}{2}v \\ v - \frac{2}{3}w \\ w \end{bmatrix}.$$

Así, x^TAx es una suma de cuadrados con los pivotes 2, $\frac{3}{2}$ y $\frac{4}{3}$ como coeficientes:

$$x^TAx = (L^Tx)^TD(L^Tx) = 2\left(u - \frac{1}{2}v\right)^2 + \frac{3}{2}\left(v - \frac{2}{3}w\right)^2 + \frac{4}{3}(w)^2.$$

Estos pivotes positivos en D multiplican cuadrados perfectos para hacer positiva a x^TAx. Así, la condición IV implica la condición I, y la demostración está completa. ■

Es hermoso que la eliminación y la completación al cuadrado sean realmente lo mismo. La eliminación retira a x_1 de todas las ecuaciones posteriores. De manera semejante, el primer

cuadrado explica todos los términos en $x^T Ax$ que implican a x_1. La suma de cuadrados tiene a los pivotes fuera. ¡*Los multiplicadores* ℓ_{ij} *están dentro*! En el ejemplo, puede ver los números $-\frac{1}{2}$ y $-\frac{2}{3}$ dentro de los cuadrados.

Todo elemento diagonal a_{ii} *debe ser positivo*. Como se sabe por los ejemplos, no obstante, es mucho más que suficiente observar sólo los elementos en la diagonal.

Los pivotes d_i no deben confundirse con los valores característicos. Para una matriz positiva definida típica, se trata de dos conjuntos completamente distintos de números positivos. En el ejemplo de 3 por 3, quizá la prueba más fácil sea la del determinante:

$$\textbf{Prueba del determinante} \quad \det A_1 = 2, \det A_2 = 3, \det A_3 = \det A = 4.$$

Los pivotes son las razones $d_1 = 2$, $d_2 = \frac{3}{2}$, $d_3 = \frac{4}{3}$. Normalmente, la prueba de los valores característicos es el cálculo más largo. Para esta A se sabe que todos los λs son positivos:

$$\textbf{Prueba del valor característico} \quad \lambda_1 = 2 - \sqrt{2}, \quad \lambda_2 = 2, \quad \lambda_3 = 2 + \sqrt{2}.$$

Aunque es la más difícil de aplicar a una simple matriz, la prueba de los valores característicos puede ser la de más utilidad para efectos teóricos. ***Cada prueba es suficiente por sí misma.***

Matrices positivas definidas y mínimos cuadrados

Espero que el lector permita la presentación de una prueba más para comprobar si una matriz es positiva definida. Dicha prueba está muy próxima. Las matrices positivas definidas se relacionaron con los pivotes (véase el capítulo 1), los determinantes (véase el capítulo 4) y con los valores característicos (véase el capítulo 5). Ahora se les ve en los problemas de mínimos cuadrados del capítulo 3, provenientes de las matrices rectangulares del capítulo 2.

La matriz rectangular será R y el problema de mínimos cuadrados será $Rx = b$. Tiene m ecuaciones con $m \geq n$ (se incluyen los sistemas cuadrados). *La elección por mínimos cuadrados* $\hat{x}$ *es la solución de* $R^T R\hat{x} = R^T b$. Esta matriz $A = R^T R$ no sólo es simétrica, sino que también es positiva definida, como se demostrará a continuación, en el supuesto de que las n columnas de R sean linealmente independientes:

6C La matriz simétrica A es positiva definida si y sólo si
(V) Existe una matriz R con columnas independientes tal que $A = R^T R$.

La clave consiste en reconocer a $x^T Ax$ *como* $x^T R^T Rx = (Rx)^T (Rx)$. Esta longitud al cuadrado $\|Rx\|^2$ es positiva (a menos que $x = 0$), porque las columnas de R son independientes. (Si x es diferente de cero, entonces Rx es diferente de cero.) Así, $x^T R^T Rx > 0$ y $R^T R$ es positiva definida.

Queda por encontrar una R para la cual $A = R^T R$. Esto ya se hizo dos veces:

$$\textbf{Eliminación} \quad A = LDL^T = (L\sqrt{D})(\sqrt{D}L^T). \quad \text{Así, tomar} \quad R = \sqrt{D}L^T.$$

Esta ***descomposición de Cholesky*** tiene separados los pivotes equitativamente entre L y L^T.

$$\textbf{Valores característicos} \quad A = Q\Lambda Q^T = (Q\sqrt{\Lambda})(\sqrt{\Lambda}Q^T). \quad \text{Así, tomar} \quad R = \sqrt{\Lambda}Q^T. \quad (3)$$

Una tercera posibilidad es $R = Q\sqrt{\Lambda}Q^T$, la ***raíz cuadrada positiva definida simétrica*** de A. Hay muchas otras opciones, cuadradas o rectangulares, que pueden verse. Si cualquier R se multiplica por *una matriz* Q *con columnas ortonormales*, entonces $(QR)^T(QR) = R^T Q^T QR = R^T IR = A$. En consecuencia, QR es otra opción.

Aplicaciones de las matrices positivas definidas se presentan en mi libro anterior *Introduction to Applied Mathematics* y también en el nuevo *Applied Mathematics and Scientific Computing* (consulte la página www.wellesleycambridge.com). Se menciona que $Ax = \lambda Mx$ se presenta constantemente en análisis ingenieril. Si A y M son positivas definidas, este problema generalizado es paralelo al conocido $Ax = \lambda x$, y $\lambda > 0$. M es una ***matriz masa*** para el *método de elementos finitos* de la sección 6.4.

Matrices semidefinidas

Las pruebas para comprobar si una matriz es positiva definida relajan $x^T Ax > 0$, $\lambda > 0$, $d > 0$, y $\det > 0$, para dejar que aparezcan ceros. La cuestión principal es ver analogías con el caso positiva definida.

6D Cada una de las pruebas siguientes constituye una prueba necesaria y suficiente para que una matriz simétrica real A sea ***positiva semidefinida***:

(I′) $x^T Ax \geq 0$ para todos los vectores x (esto define a las matrices positivas semidefinidas).
(II′) Todos los valores característicos de A satisfacen $\lambda_i \geq 0$.
(III′) Ninguna submatriz principal tienen determinantes negativos.
(IV′) Ningunos pivotes son negativos.
(V′) Existe una matriz R, quizá con columnas dependientes, tal que $A = R^T R$.

La diagonalización de $A = Q\Lambda Q^T$ conduce a $x^T Ax = x^T Q\Lambda Q^T x = y^T \Lambda y$. Si el rango de A es r, en $y^T \Lambda y = \lambda_1 y_1^2 + \cdots + \lambda_r y_r^2$ hay r λs diferentes de cero y r cuadrados perfectos.

Nota La novedad es que la condición III′ es válida para todas las submatrices principales, no sólo para aquellas que están en la esquina superior izquierda. En caso contrario, no sería posible distinguir entre dos matrices cuyos determinantes superiores izquierdos fuesen todos cero:

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \text{ es positiva semidefinida, y } \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} \text{ es negativa semidefinida.}$$

Para mantener la simetría, un intercambio de renglones va acompañado de un intercambio de columnas.

Ejemplo 2

$$A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \text{ es positiva } \textit{semi}\text{definida, por todas las cinco pruebas:}$$

(I′) $x^T Ax = (x_1 - x_2)^2 + (x_1 - x_3)^2 + (x_2 - x_3)^2 \geq 0$ (cero si $x_1 = x_2 = x_3$).
(II′) Los valores característicos son $\lambda_1 = 0$, $\lambda_2 = \lambda_3 = 3$ (y cero los valores característicos).
(III′) $\det A = 0$ y los determinantes menores son positivos.

$$\text{(IV′)} \quad A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \to \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{3}{2} & -\frac{3}{2} \\ 0 & -\frac{3}{2} & \frac{3}{2} \end{bmatrix} \to \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & 0 & \mathbf{0} \end{bmatrix} \text{ (pivote faltante).}$$

(V′) $A = R^T R$ con columnas dependientes en R:

$$\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \quad (1, 1, 1) \text{ en el espacio nulo.}$$

Observación Las condiciones para que una matriz sea semidefinida también pudieron deducirse de las condiciones originales I a V si se usa el siguiente truco: sumar un pequeño múltiplo de la identidad para obtener una matriz positiva definida $A + \epsilon I$. Luego, ϵ tiende a cero. Debido a que los determinantes y los valores característicos dependen continuamente de $\epsilon = 0$, son positivos hasta el último momento. En ϵ deben seguir siendo no negativos.

A menudo, mis alumnos me preguntan sobre matrices positiva definida *no simétricas*. Nunca se utiliza este término. Una definición razonable es que la parte simétrica $\frac{1}{2}(A + A^T)$ debe ser positiva definida. Eso garantiza que *las partes reales de los valores característicos son positivas.* Aunque esto no es necesario: $A = \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}$ tiene $\lambda > 0$ pero $\frac{1}{2}(A + A^T) = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ es indefinida.

$$\text{Si } Ax = \lambda x, \text{ entonces } x^H Ax = \lambda x^H x \text{ y } x^H A^H x = \bar{\lambda} x^H x.$$
$$\text{Al sumar } \tfrac{1}{2} x^H(A + A^H)x = (\text{Re}\,\lambda)x^H x > 0 \quad \text{de modo que Re}\lambda > 0.$$

Elipsoides en *n* dimensiones

En todo este libro, la geometría ha sido de ayuda para el álgebra de matrices. Una ecuación lineal produjo un plano. El sistema $Ax = b$ proporciona una intersección de planos. Los mínimos cuadrados proporcionaron una proyección perpendicular. El determinante es el volumen de una caja. Luego, para una matriz positiva definida y su $x^T Ax$, finalmente se obtiene una figura curva. Se trata de una *elipse* en dos dimensiones, y de un *elipsoide* en *n* dimensiones.

La ecuación que debe considerarse es $x^T Ax = 1$. Si A es la matriz identidad, esto se simplifica a $x_1^2 + x_2^2 + \cdots + x_n^2 = 1$. Ésta es la ecuación de la "esfera unitaria" en $\mathbf{R}^n$. Si $A = 4I$, la esfera se vuelve más chica. La ecuación cambia a $4x_1^2 + \cdots + 4x_n^2 = 1$. En vez de pasar por $(1, 0, \ldots, 0)$, lo hace por $(\frac{1}{2}, 0, \ldots, 0)$. El centro está en el origen, ya que si x satisface $x^T Ax = 1$, también lo hace el vector opuesto $-x$. El paso importante es pasar de la matriz identidad a una *matriz diagonal*:

$$\textbf{Elipsoide} \quad \text{Para } A = \begin{bmatrix} 4 & & \\ & 1 & \\ & & \frac{1}{9} \end{bmatrix}, \text{ la ecuación es } x^T Ax = 4x_1^2 + x_2^2 + \tfrac{1}{9}x_3^2 = 1.$$

Debido a que los elementos son desiguales (¡y positivos!), la esfera cambia a un elipsoide.

Una solución es $x = (\frac{1}{2}, 0, 0)$ a lo largo del primer eje. Otra es $x = (0, 1, 0)$. El eje mayor tiene el punto más alejado $x = (0, 0, 3)$. Se parece a una pelota de fútbol americano o rugby, aunque no exactamente; éstas están más próximas a $x_1^2 + x_2^2 + \frac{1}{2}x_3^2 = 1$. Los dos coeficientes iguales las hacen circulares en el plano x_1-x_2, ¡y más fáciles de lanzar!

A continuación se presenta el paso final, para permitir elementos diferentes de cero lejos de la diagonal de A.

Ejemplo 3 $A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$ y $x^T Ax = 5u^2 + 8uv + 5v^2 = 1$. Esta elipse está centrada en $u = v = 0$, aunque los ejes no son tan claros. Los 4s fuera de la diagonal dejan positiva definida a la matriz, aunque hacen rotar la elipse, ya que sus ejes dejan de estar alineados con los ejes de coordenadas (véase la figura 6.2). Se demostrará que ***los ejes de las elipses apuntan hacia los vectores característicos de A.*** Debido a que $A = A^T$, estos vectores característicos y los ejes son ortogonales. El eje *mayor* de la elipse corresponde al valor característico *más pequeño* de A.

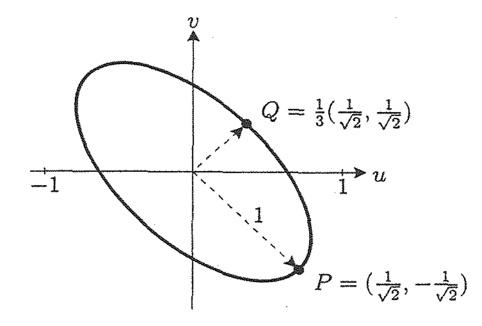


Figura 6.2 La elipse $x^T Ax = 5u^2 + 8uv + 5v^2 = 1$ y sus ejes principales.

Para localizar la elipse se calculan $\lambda_1 = 1$ y $\lambda_2 = 9$. Los vectores característicos unitarios son $(1, -1)/\sqrt{2}$ y $(1, 1)/\sqrt{2}$. Estos son ángulos de 45° con los ejes *u*-*v*, y están alineados con los ejes de la elipse. La forma para ver correctamente a la elipse es *volver a escribir* $x^T Ax = 1$:

$$\textbf{Nuevos cuadrados} \quad 5u^2 + 8uv + v^2 = \left(\frac{u}{\sqrt{2}} - \frac{v}{\sqrt{2}}\right)^2 + 9\left(\frac{u}{\sqrt{2}} + \frac{v}{\sqrt{2}}\right)^2 = 1. \tag{4}$$

$\lambda = 1$ y $\lambda = 9$ están fuera de los cuadrados. Los vectores característicos están dentro. Esto es diferente para completar el cuadrado a $5\left(u + \frac{4}{5}v\right)^2 + \frac{9}{5}v^2$, con los *pivotes* fuera.

El primer cuadrado es igual a 1 en $(1/\sqrt{2}, -1/\sqrt{2})$ al final del eje mayor. El eje menor es un tercio más corto, ya que para cancelar el 9 se requiere $\left(\frac{1}{3}\right)^2$.

Cualquier elipsoide $x^T Ax = 1$ puede simplificarse de la misma manera. *El paso clave es diagonalizar* $A = Q\Lambda Q^T$. La figura se enderezó al hacer girar los ejes. Algebraicamente, el cambio a $y = Q^T x$ produce una suma de cuadrados:

$$x^T Ax = (x^T Q)\Lambda(Q^T x) = y^T \Lambda y = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 = 1. \tag{5}$$

El *eje mayor* tiene $y_1 = 1/\sqrt{\lambda_1}$ a lo largo del vector característico con el menor valor característico.

Los otros ejes están en dirección de los otros vectores característicos. Sus longitudes son $1/\sqrt{\lambda_2}, \ldots, 1/\sqrt{\lambda_n}$. Observe que los λs deben ser positivos *–la matriz debe ser positiva definida–* o estas raíces cuadradas están en problemas. Una ecuación indefinida $y_1^2 - 9y_2^2 = 1$ describe una hipérbola y no una elipse. Una hipérbola es una sección transversal que pasa por una silla, y una elipse es una sección transversal que pasa por un tazón.

El cambio de x a $y = Q^T x$ hace girar los ejes del espacio para coincidir con los ejes del elipsoide. En las variables y puede verse que se trata de un elipsoide, ya que la ecuación se vuelve manipulable:

6E Suponga que $A = Q\Lambda Q^T$ con $\lambda_i > 0$. Al hacer girar $y = Q^T x$ se simplifica $x^T Ax = 1$:

$$x^T Q\Lambda Q^T x = 1, \quad y^T \Lambda y = 1, \quad \text{y} \quad \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 = 1.$$

Esta es la ecuación de un elipsoide. Sus ejes tienen longitudes $1/\sqrt{\lambda_1}, \ldots, 1/\sqrt{\lambda_n}$ a partir del centro. En el espacio x original apuntan a lo largo de los vectores característicos de A.

La ley de la inercia

Para eliminación y valores característicos, las matrices se vuelven más sencillas mediante operaciones elementales. La cuestión esencial es saber cuáles propiedades de una matriz permanecen sin cambio. Cuando un múltiplo de un renglón se resta de otro renglón, el espacio renglón, el espacio nulo, el rango y el determinante –todos– permanecen igual. Para valores característicos, la operación básica fue una transformación de semejanza $A \to S^{-1}AS$ (o $A \to M^{-1}AM$). Los valores característicos permanecen sin cambio (así como la forma de Jordan). Ahora se pregunta lo mismo para matrices simétricas: *¿cuáles son las operaciones elementales y sus invariantes para* x^TAx?

La operación básica sobre una forma cuadrática es cambiar variables. Un nuevo vector y está relacionado con x por medio de alguna matriz no singular, $x = Cy$. La forma cuadrática se convierte en y^TC^TACy. Esto muestra la operación fundamental sobre A:

$$\textbf{Transformación de congruencia} \quad A \to C^TAC \quad \text{para alguna } C \text{ no singular.} \qquad (6)$$

La simetría de A se preserva, ya que C^TAC permanece simétrica. La pregunta verdadera es ¿qué otras propiedades comparten A y C^TAC? La respuesta la proporciona la ***ley de la inercia*** de Sylvester.

6F C^TAC tiene el mismo número de valores característicos positivos, valores característicos negativos, y valores característicos cero que A.

Los *signos* de los valores característicos (y no los valores característicos mismos) se preservan con una transformación de congruencia. En la demostración, se supondrá que A es no singular. Así, C^TAC también es no singular, y no hay valores característicos de que preocuparse. (En caso contrario, es posible trabajar con la no singular $A + \epsilon I$ y $A - \epsilon I$, y al final hacer $\epsilon \to 0$.)

Demostración Tomaremos prestado un truco de topología. Suponga que C está relacionada con una matriz ortogonal Q, por medio de una cadena continua de matrices no singulares $C(t)$. En $t = 0$ y $t = 1$, $C(0) = C$ y $C(1) = Q$. Así, los valores característicos de $C(t)^TAC(t)$ cambian gradualmente, cuando t va de 0 a 1, de los valores característicos de C^TAC a los valores característicos de Q^TAQ. Debido a que $C(t)$ nunca es singular, *ninguno de estos valores característicos puede tocar a cero* (¡por no decir cruzarlo!) En consecuencia, el número de valores característicos a la derecha de cero, y el número a la izquierda, es el mismo para C^TAC que para Q^TAQ. Y A tiene exactamente los mismos valores característicos que la matriz semejante $Q^{-1}AQ = Q^TAQ$.

Una buena elección de Q consiste en aplicar el proceso de Gram-Schmidt a las columnas de C. Así, $C = QR$, y la cadena de matrices es $C(t) = tQ + (1-t)QR$. La familia $C(t)$ va lentamente a través del proceso de Gram-Schmidt, de QR a Q. Es invertible porque Q es invertible y la diagonal del factor triangular $tI + (1-t)R$ es positiva. Así se termina la demostración. ■

Ejemplo 4 Suponga que $A = I$. Entonces $C^TAC = C^TC$ es positiva definida. Tanto I como C^TC tienen n valores característicos positivos, lo cual confirma la ley de la inercia.

Ejemplo 5 Si $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, entonces C^TAC tiene un determinante negativo:

$$\det C^TAC = (\det C^T)(\det A)(\det C) = -(\det C)^2 < 0.$$

Entonces C^TAC debe tener un valor característico positivo y un valor característico negativo, como A.

Ejemplo 6 Esta es la aplicación importante:

6G Para cualquier matriz simétrica A, ***los signos de los pivotes coinciden con los signos de los valores característicos.*** La matriz valor característico Λ y la matriz pivote D tienen el mismo número de elementos positivos, elementos negativos, y elementos cero.

Se supondrá que A permite la factorización simétrica $A = LDL^T$ (sin intercambios de renglones). Por la ley de la inercia, A tiene el mismo número de valores característicos positivos que D. Sin embargo, los valores característicos de D son justo sus elementos diagonales (los pivotes). Así, el número de pivotes positivos coincide con el número de valores característicos positivos de A.

Lo anterior es hermoso y práctico. Es hermoso porque conlleva (para matrices simétricas) dos partes de este libro que previamente estaban separadas: los *pivotes* y los *valores característicos*. También es práctico, ya que los pivotes pueden localizar a los valores característicos:

A tiene pivotes positivos
$A - 2I$ tiene un pivote negativo

$$A = \begin{bmatrix} 3 & 3 & 0 \\ 3 & 10 & 7 \\ 0 & 7 & 8 \end{bmatrix} \qquad A - 2I = \begin{bmatrix} 1 & 3 & 0 \\ 3 & 8 & 7 \\ 0 & 7 & 6 \end{bmatrix}.$$

Gracias a nuestra prueba, los valores característicos de A son positivos. Pero se sabe que $\lambda_{\text{mín}}$ es *menor que* 2, porque al restar 2 hizo que éste cayera bajo cero. En el siguiente paso se aborda $A - I$, para ver $\lambda_{\text{mín}} < 1$. (Esto se debe a que $A - I$ tiene un pivote negativo). El intervalo que contiene a λ se divide a la mitad en cada paso al comprobar el signo de los pivotes.

Este era casi el primer método práctico para calcular valores característicos. Era dominante alrededor de 1960, después de una mejora importante: hacer tridiagonal a A primero. Luego, los pivotes se calculaban en $2n$ pasos, en vez de en $\frac{1}{6}n^3$. La eliminación se vuelve rápida, y la búsqueda de valores característicos (al dividir en dos los intervalos) se vuelve simple. El favorito actual es el método QR que se presenta en el capítulo 7.

El problema generalizado de valores característicos

En sus problemas de valores característicos, la física, la ingeniería, y la estadística suelen ser suficientemente bondadosas para producir matrices simétricas. ***Sin embargo, algunas veces $Ax = \lambda x$ se sustituye por $Ax = \lambda Mx$.*** *Hay dos matrices, en vez de una.*

Un ejemplo es el movimiento de dos masas distintas en una línea de resortes:

$$\begin{aligned} m_1\frac{d^2v}{dt^2} + 2v - w &= 0 \\ m_2\frac{d^2w}{dt^2} - v + 2w &= 0 \end{aligned} \quad \text{o bien,} \quad \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}\frac{d^2u}{dt^2} + \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}u = 0. \qquad (7)$$

Cuando las masas son iguales, $m_1 = m_2 = 1$, este era el sistema anterior $u'' + Au = 0$. Ahora es $Mu'' + Au = 0$, que es una *matriz "masa"* M. El problema de valores característicos surge cuando se buscan soluciones exponenciales $e^{i\omega t}x$:

$$Mu'' + Au = 0 \quad \text{se convierte en} \quad M(i\omega)^2e^{i\omega t}x + Ae^{i\omega t}x = 0. \qquad (8)$$

Al cancelar $e^{i\omega t}$, y al escribir λ por ω^2, este es un problema de valores característicos:

$$\textbf{Problema generalizado} \quad Ax = \lambda Mx \qquad \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}x = \lambda\begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}x. \qquad (9)$$

Existe una solución cuando $A - \lambda M$ es singular. La elección especial $M = I$ devuelve la expresión de costumbre $\det(A - \lambda I) = 0$. $\det(A - \lambda M)$ se trabajará con $m_1 = 1$ y $m_2 = 2$:

$$\det\begin{bmatrix} 2-\lambda & -1 \\ -1 & 2-2\lambda \end{bmatrix} = 2\lambda^2 - 6\lambda + 3 = 0 \quad \text{proporciona} \quad \lambda = \frac{3 \pm \sqrt{3}}{2}.$$

Para el vector característico $x_1 = (\sqrt{3} - 1, 1)$, las dos masas oscilan juntas, aunque la primera sólo se mueve $\sqrt{3} - 1 \approx 0.73$ En el modo más rápido, las componentes de $x_2 = (1 + \sqrt{3} - 1)$ tienen signos opuestos y las masas se mueven en direcciones opuestas. Esta vez la masa más pequeña avanza mucho más.

La teoría subyacente es más fácil de explicar si M se separa en R^TR. (*Se supone que M es positiva definida*). Luego, la sustitución $y = Rx$ cambia

$$Ax = \lambda Mx = \lambda R^TRx \quad \text{hacia} \quad AR^{-1}y = \lambda R^Ty.$$

Al escribir C por R^{-1}, y multiplicar por $(R^T)^{-1} = C^T$, lo anterior se convierte en un problema de valores característicos estándar para la matriz simétrica *simple* C^TAC:

$$\textbf{Problema equivalente} \qquad C^TACy = \lambda y. \tag{10}$$

Los valores característicos λ_j son los mismos que para la $Ax = \lambda Mx$ original, y los vectores característicos están relacionados por $y_j = Rx_j$. Las propiedades de C^TAC conducen directamente a las propiedades de $Ax = \lambda Mx$, cuando $A = A^T$ y M es positiva definida:

1. Los valores característicos para $Ax = \lambda Mx$ son reales, porque C^TAC es simétrica.
2. Por la ley de la inercia, los λs tienen el mismo signo que los valores característicos de A.
3. C^TAC tiene vectores característicos ortogonales y_j. Así, los vectores característicos de $Ax = \lambda Mx$ tienen

$$\textbf{“}M\textbf{-ortogonalidad”} \qquad x_i^TMx_j = x_i^TR^TRx_j = y_i^Ty_j = 0. \tag{11}$$

A y M se *diagonalizan simultáneamente*. Si S tiene los x_j en sus columnas, entonces $S^TAS = \Lambda$ y $S^TMS = I$. Esta es una transformación de *congruencia*, con S^T en la izquierda, y no una transformación de semejanza con S^{-1}. Es fácil resumir la cuestión importante: En tanto M es positiva definida, el problema generalizado de valores característicos $Ax = \lambda Mx$ se comporta exactamente como $Ax = \lambda x$.

Conjunto de problemas 6.2

1. ¿Para qué intervalo de números a y b ocurre que las matrices A y B son positiva definida?

$$A = \begin{bmatrix} a & 2 & 2 \\ 2 & a & 2 \\ 2 & 2 & a \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 & 4 \\ 2 & b & 8 \\ 4 & 8 & 7 \end{bmatrix}.$$

2. Decida si las siguientes matrices son positiva definida o no:

$$A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}^2.$$

3. Construya una *matriz indefinida* cuyos elementos más grandes estén en la diagonal principal:

$$A = \begin{bmatrix} 1 & b & -b \\ b & 1 & b \\ -b & b & 1 \end{bmatrix} \quad \text{con } |b| < 1 \text{ puede tener } \det A < 0.$$

4. Demuestre a partir de los valores característicos que si A es positiva definida, entonces también lo son A^2 y A^{-1}.

5 *Si A y B son positiva definida, entonces $A + B$ es positiva definida*. Los pivotes y los valores característicos no son convenientes para $A + B$. Es mucho mejor intentar $x^T(A + B)x > 0$.

6. A partir de los pivotes, los valores característicos y los vectores característicos de $A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$, escriba A como R^TR en tres formas:

$$(L\sqrt{D})(\sqrt{D}L^T), (Q\sqrt{\Lambda})\ (\sqrt{\Lambda}Q^T), \ \text{y} \ (Q\sqrt{\Lambda}Q^T)(Q\sqrt{\Lambda}Q^T).$$

7. Si $A = Q\Lambda Q^T$, es positiva definida simétrica, entonces $R = Q\sqrt{\Lambda}Q^T$ es su *raíz cuadrada positiva definida simétrica*. ¿Por qué los valores característicos de R son positivos? Calcule R, y compruebe que $R^2 = A$ para

$$A = \begin{bmatrix} 10 & 6 \\ 6 & 10 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 10 & -6 \\ -6 & 10 \end{bmatrix}.$$

8. Si A, es positiva definida simétrica y C es no singular, demuestre que $B = C^TAC$ también es positiva definida simétrica.

9. Si $A = R^TR$, demuestre la desigualdad de Schwarz generalizada $|x^TAy|^2 \le (x^TAx)(y^TAy)$.

10. La elipse $u^2 + 4v^2 = 1$ corresponde a $A = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$. Escriba los valores característicos y los vectores característicos, y trace la elipse.

11. Reduzca la ecuación $3u^2 - 2\sqrt{2}uv + 2v^2 = 1$. a una suma de cuadrados encontrando los valores característicos de la A correspondiente, y trace la elipse.

12. En tres dimensiones, $\lambda_1y_1^2 + \lambda_2y_2^2 + \lambda_3y_3^2 = 1$ representa un elipsoide cuando todos los $\lambda_i > 0$. Describa todas las distintas clases de superficies que aparecen en el caso positiva semidefinida cuando uno o más de los valores característicos es (son) cero.

13. Escriba las cinco condiciones para que una matriz de 3 por 3 sea *negativa definida* ($-A$ es positiva definida) con atención especial a la condición III: ¿Cómo está relacionado $\det(-A)$ con $\det A$?

14. Decida si las siguientes matrices son positivas definidas, negativas definidas, semidefinidas, o indefinidas:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 4 \\ 3 & 4 & 9 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 6 & -2 & 0 \\ 0 & -2 & 5 & -2 \\ 0 & 0 & -2 & 3 \end{bmatrix}, \quad C = -B, \quad D = A^{-1}.$$

¿Existe una solución real para $-x^2 - 5y^2 - 9z^2 - 4xy - 6xz - 8yz = 1$?

15. Suponga que A es positiva definida simétrica y que Q es una matriz ortogonal. ¿Falso o verdadero?
a) Q^TAQ es una matriz diagonal.

b) Q^TAQ es positiva definida simétrica.
c) Q^TAQ tiene los mismos valores característicos que A.
d) e^{-A} es positiva definida simétrica.

16. Si A es positiva definida y se incrementa a_{11}, a partir de cofactores demuestre que el determinante se ha incrementado. Demuestre con un ejemplo que lo anterior puede fallar si A es indefinida.

17. A partir de $A = R^TR$, demuestre para matrices positivas definidas que $\det A \leq a_{11}a_{22}\cdots a_{nn}$. (La longitud al cuadrado de la columna j de R es a_{jj}. Use determinante = volumen.)

18. (Prueba de M de estabilidad de Lyapunov) Suponga que $AM + M^HA = -I$ con A positiva definida. Si $Mx = \lambda x$, demuestre que $\operatorname{Re}\lambda < 0$. (*Sugerencia*: Multiplique la primera ecuación por x^H y x.)

19. ¿Cuáles matrices simétricas A de 3 por 3 producen las siguientes funciones $f = x^TAx$? ¿Por qué la primera matriz es positiva definida pero la segunda no?

a) $f = 2(x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_2x_3)$.
b) $f = 2(x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_1x_3 - x_2x_3)$.

20. Calcule los tres determinantes superiores izquierdos para establecer la condición de positiva definida. Compruebe que sus razones proporcionan los pivotes segundo y tercero.

$$A = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 5 & 3 \\ 0 & 3 & 8 \end{bmatrix}.$$

21. Una matriz positiva definida no puede tener un cero (o incluso peor: un número negativo) en su diagonal. Demuestre que esta matriz no cumple $x^TAx > 0$:

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 4 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{no es positiva cuando} \quad (x_1, x_2, x_3) = (\quad,\quad,\quad).$$

22. Un elemento diagonal a_{jj} de una matriz simétrica no puede ser menor que todos los λs. Si lo fuese, entonces $A - a_{jj}I$ tendría _____ valores característicos y posiblemente sería positiva definida. Sin embargo, $A - a_{jj}I$ tiene un _____ en la diagonal principal.

23. Proporcione una razón rápida de por qué las siguientes afirmaciones son verdaderas:

a) Toda matriz positiva definida es invertible.
b) La única matriz proyección positiva definida es $P = I$.
c) Una matriz diagonal con elementos diagonales positivos es positiva definida.
d) ¡Una matriz simétrica con un determinante positivo podría no ser positiva definida!

24. ¿Para cuáles s y t se cumple que A y B tienen a todos los $\lambda > 0$ (y por tanto son positivas definidas)?

$$A = \begin{bmatrix} s & -4 & -4 \\ -4 & s & -4 \\ -4 & -4 & s \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} t & 3 & 0 \\ 3 & t & 4 \\ 0 & 4 & t \end{bmatrix}.$$

25. Tal vez usted ya haya visto la ecuación de una elipse como $\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1$. ¿Cuáles son a y b cuando la ecuación se escribe como $\lambda_1x^2 + \lambda_2y^2 = 1$? La elipse $9x^2 + 16y^2 = 1$ tiene semiejes cuyas longitudes son $a =$ _____ y $b =$ _____.

26. Trace la elipse inclinada $x^2 + xy + y^2 = 1$ y encuentre las semilongitudes de sus ejes, a partir de los valores característicos de la A correspondiente.

27. Con pivotes positivos en D, la factorización $A = LDL^T$ se convierte en $L\sqrt{D}\sqrt{D}L^T$. (Las raíces cuadradas de los pivotes proporcionan $D = \sqrt{D}\sqrt{D}$. Así, $C = L\sqrt{D}$ conduce a la ***factorización de Cholesky*** $A = CC^T$, que es "simetrizada LU":

$$\text{A partir de } C = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \text{ encuentre } A. \quad \text{A partir de } A = \begin{bmatrix} 4 & 8 \\ 8 & 25 \end{bmatrix} \text{ encuentre } C.$$

28. En la factorización de Cholesky $A = CC^T$, con $C = L\sqrt{D}$, las raíces cuadradas de los pivotes están en la diagonal de C. Encuentre C (triangular inferior) para

$$A = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 8 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 7 \end{bmatrix}.$$

29. La factorización simétrica $A = LDL^T$ significa que $x^TAx = x^TLDL^Tx$:

$$\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 1 & 0 \\ b/a & 1 \end{bmatrix}\begin{bmatrix} a & 0 \\ 0 & (ac-b^2)/a \end{bmatrix}\begin{bmatrix} 1 & b/a \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}.$$

El miembro izquierdo es $ax^2 + 2bxy + cy^2$. El miembro derecho es $a\left(x + \frac{b}{a}y\right)^2 +$ _____y^2.

¡El segundo pivote completa el cuadrado! Pruebe con $a = 2$, $b = 4$, $c = 10$.

30. Sin multiplicar $A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}\begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}$, encuentre:

a) El determinante de A.
b) Los valores característicos de A.
c) Los vectores característicos de A.
d) Una razón de por qué A es positiva definida simétrica.

31. Para las matrices semidefinidas

$$A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \text{ (rango 2)} \quad \text{y} \quad B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \text{ (rango 1)},$$

Escriba x^TAx como una suma de dos cuadrados y x^TBx como un cuadrado.

32. Aplique tres pruebas cualesquiera para probar cada una de las siguientes matrices

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & 1 & 2 \end{bmatrix},$$

y decidir si son positivas definidas, positivas semidefinidas, o indefinidas.

33. Para $C = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$ y $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, confirme que C^TAC tiene valores característicos de los mismos signos que A. Construya una cadena de matrices no singulares $C(t)$ que relacione a C con una Q ortogonal. ¿Por qué es imposible construir una cadena no singular que relacione a C con la matriz identidad?

34. Si todos los pivotes de una matriz son mayores que 1, ¿es cierto que todos los valores característicos son mayores que 1? Pruebe lo anterior con las matrices tridiagonales $-1, 2, -1$.

35. Use los pivotes de $A - \frac{1}{2}I$ para decidir si A tiene un valor característico menor que $\frac{1}{2}$:

$$A - \frac{1}{2}I = \begin{bmatrix} 2.5 & 3 & 0 \\ 3 & 9.5 & 7 \\ 0 & 7 & 7.5 \end{bmatrix}.$$

36. Una demostración algebraica de la *ley de la inercia* empieza con los vectores característicos ortonormales $x_1, \ldots, x_p$ de A correspondientes a los valores característicos $\lambda_i > 0$, y los vectores característicos ortonormales $y_1, \ldots, y_q$ de C^TAC correspondientes a los valores característicos $\mu_i < 0$.

a) Para demostrar que los $p + q$ vectores $x_1, \ldots, x_p, Cy_1, \ldots, Cy_q$ son independientes, suponga que alguna combinación proporciona cero:

$$a_1x_1 + \cdots + a_px_p = b_1Cy_1 + \cdots + b_qCy_q \ (= z, \text{ digamos}).$$

Demuestre que $z^TAz = \lambda_1a_1^2 + \cdots + \lambda_pa_p^2 \geq 0$ y $z^TAz = \mu_1b_1^2 + \cdots + \mu_qb_q^2 \leq 0$.

b) Deduzca que las *as* y las *bs* son cero (demostrando independencia lineal). Con base en lo anterior, deduzca que $p + q \leq n$.

c) Con el mismo razonamiento para los $n - p$ λs negativos y los $n - q$ μs positivos se obtiene $n - p + n - q \leq n$. (Nuevamente se supone que no hay valores característicos diferentes de cero, que se manejan por separado). Demuestre que $p + q = n$, de modo que el número p de λs positivos es igual al número $n - q$ de μs positivos; lo cual es la ley de la inercia.

37. Si C es no singular, demuestre que A y C^TAC tienen el mismo rango. Así, tienen el mismo número de valores característicos cero.

38. Por experimentación, encuentre el número de valores característicos positivos, negativos, y cero de

$$A = \begin{bmatrix} I & B \\ B^T & 0 \end{bmatrix}$$

cuando el bloque B (de orden $\frac{1}{2}n$) es no singular.

39. ¿A y C^TAC siempre satisfacen la ley de la inercia cuando C no es cuadrada?

40. En la ecuación (9) con $m_1 = 1$ y $m_2 = 2$, compruebe que los modos normales son M ortogonales: $x_1^TMx_2 = 0$.

41. Encuentre los valores característicos y los vectores característicos de $Ax = \lambda Mx$:

$$\begin{bmatrix} 6 & -3 \\ -3 & 6 \end{bmatrix} x = \frac{\lambda}{18} \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix} x.$$

42. Si las matrices simétricas A y M son indefinidas, $Ax = \lambda Mx$ podría no tener valores característicos reales. Construya un ejemplo de 2 por 2.

43. Un *grupo* de matrices singulares incluye $A\,B$ y A^{-1} si incluye a A y B. "Los productos e inversas permanecen en el grupo". ¿Cuáles de los siguientes conjuntos son grupos? *Matrices positivas definidas simétricas A, matrices ortogonales Q, todas las exponenciales e^{tA} de una matriz fija A, las matrices P con valores característicos positivos, matrices D con determinante* 1. Invente un grupo que sólo contenga matrices positivas definidas.

6.3 DESCOMPOSICIÓN DEL VALOR SINGULAR

Para el final del curso básico se ha dejado una gran factorización matricial. $U\,\Sigma V^T$ se vincula con LU de la eliminación y con QR de la ortogonalización (Gauss y proceso de Gram-Schmidt). No se le ha asignado ninguna denominación o nombre: $A = U\,\Sigma V^T$ se conoce como "DVS" o ***descomposición del valor singular***. Así como "SVD" por sus siglas en inglés. Se quiere describirlo, demostrarlo, y analizar sus aplicaciones, que son muchas y cada vez más numerosas.

La DVS está estrechamente relacionada con la factorización valores característicos-vectores característicos $Q\Lambda Q^T$ de una matriz positiva definida. Los valores característicos están en la diagonal de la matriz Λ. La matriz vector característico Q *es ortogonal* ($Q^TQ = I$) porque es posible escoger que los vectores característicos de una matriz simétrica sean ortonormales. Esto no es cierto para la mayor parte de las matrices, y para las matrices rectangulares es ridículo (porque los valores característicos están indefinidos). Sin embargo, ahora se permite que Q a la izquierda y Q^T a la derecha sean *dos matrices ortogonales cualesquiera* U y V^T, no necesariamente traspuestas entre sí. Así, toda matriz puede separarse en $A = U\Sigma V^T$.

¡La matriz diagonal (pero rectangular) Σ tiene valores característicos de A^TA, no de A! Estos elementos positivos (que también se denominan sigma) son $\Sigma_1, \ldots, \Sigma_r$. Son los ***valores singulares*** de A, y ocupan los r primeros sitios sobre la diagonal principal de Σ, cuando el rango de A es r. El resto de Σ es cero.

Con matrices rectangulares, casi siempre la clave es considerar A^TA y AA^T.

Descomposición del valor singular: Cualquier matriz A de m por n puede factorizarse como

$$A = U\Sigma V^T = \textbf{(ortogonal)(diagonal)(ortogonal)}.$$

Las columnas de U (que es de m por m) son vectores característicos de AA^T, y las columnas de V (que es de n por n) son los vectores característicos de A^TA. Los r valores singulares en la diagonal de Σ (m por n) son las raíces cuadradas de los valores característicos diferentes de cero tanto de AA^T como de A^TA.

Observación 1 Para matrices positivas definidas, Σ es Λ y $U\,\Sigma V^T$ es idéntica a $Q\Lambda Q^T$. Para otras matrices simétricas, cualesquiera valores característicos negativos en Λ se vuelven positivos en Σ. Para matrices complejas, Σ permanece real pero U y V se convierten en *unitarias* (la versión compleja de las matrices ortogonales). Se toman conjugados complejos en $U^HU = I$ y $V^HV = I$ y $A = U\,\Sigma V^H$.

Observación 2 U y V constituyen bases ortonormales de *todos los cuatro subespacios fundamentales*:

Las primeras	r	columnas de U:	**espacio columna** de A
Las últimas	$m - r$	columnas de U:	**espacio nulo izquierdo** de A
Las primeras	r	columnas de V:	**espacio renglón** de A
Las últimas	$n - r$	columnas de V:	**espacio nulo de** A

Observación 3 La DVS escoge estas bases de una manera extremadamente especial. Son más que simplemente ortonormales. *Cuando A multiplica a una columna v_j de V, produce Σ_j veces una columna de U.* Este hecho proviene directamente de $AV = U\Sigma$, considerada una columna a la vez.

Observación 4 Los vectores característicos de AA^T y A^TA deben ir en las columnas de U y V:

$$AA^T = (U\Sigma V^T)(V\Sigma^T U^T) = U\Sigma\Sigma^T U^T \quad \text{similar, y} \quad A^TA = V\Sigma^T\Sigma V^T. \quad (1)$$

U debe ser la matriz vector característico para AA^T. La matriz valor característico que está en medio es $\Sigma\Sigma^T$, que es de m por m con $\sigma_1^2, \ldots, \sigma_r^2$ en la diagonal.

Con base en $A^TA = V\Sigma^T\Sigma V^T$, la matriz *$V$ debe ser la matriz vector característico para A^TA*. La matriz diagonal $\Sigma^T\Sigma$ tiene los mismos $\sigma_1^2, \ldots, \sigma_r^2$, pero es de n por n.

Observación 5 Esta es la razón por la que $Av_j = \sigma_j u_j$. Se empieza con $A^TAv_j = \sigma_j^2 Av_j$:

Multiplicar por A $$AA^TAv_j = \sigma_j^2 Av_j \quad (2)$$

¡Lo anterior indica que Av_j es un vector característico de AA^T! Simplemente se movieron los paréntesis a $(AA^T)(Av_j)$. La longitud de este vector característico Av_j es Σ_j, ya que

$$v^TA^TAv_j = \sigma_j^2 v_j^T v_j \quad \text{proporciona} \quad \|Av_j\|^2 = \sigma_j^2.$$

De modo que el vector característico unitario es $Av_j/\sigma_j = u_j$. **En otras palabras, $AV = U\Sigma$.**

Ejemplo 1 Esta A sólo tiene una columna: rango $r = 1$. Entonces, Σ sólo tiene $\Sigma_1 = 3$:

$$\textbf{DVS} \quad A = \begin{bmatrix}-1\\2\\2\end{bmatrix} = \begin{bmatrix}-\frac{1}{3} & \frac{2}{3} & \frac{2}{3}\\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3}\\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3}\end{bmatrix}\begin{bmatrix}3\\0\\0\end{bmatrix}[\,1\,] = U_{3\times3}\,\Sigma_{3\times1}\,V_{1\times1}^T.$$

A^TA es de 1 por 1, mientras AA^T es de 3 por 3. Ambas tienen valor característico 9 (cuya raíz cuadrada es 3 en Σ). Los dos valores característicos cero de AA^T dejan algo de libertad para los vectores característicos en las columnas 2 y 3 de U. Esta matriz debe mantenerse ortogonal.

Ejemplo 2 Ahora el rango de A es 2, y $AA^T = \begin{bmatrix}2 & -1\\-1 & 2\end{bmatrix}$ con $\lambda = 3$ y 1:

$$\begin{bmatrix}-1 & 1 & 0\\0 & -1 & 1\end{bmatrix} = U\Sigma V^T = \frac{1}{\sqrt{2}}\begin{bmatrix}-1 & 1\\1 & 1\end{bmatrix}\begin{bmatrix}\sqrt{3} & 0 & 0\\0 & 1 & 0\end{bmatrix}\begin{bmatrix}1 & -2 & 1\\-1 & 0 & 1\\1 & 1 & 1\end{bmatrix}\begin{matrix}/\sqrt{6}\\/\sqrt{2}\\/\sqrt{3}\end{matrix}.$$

Observe a $\sqrt{3}$ y $\sqrt{1}$. Las columnas de U son vectores singulares *izquierdos* (vectores característicos unitarios de AA^T). Las columnas de V son vectores singulares *derechos* (vectores característicos unitarios de A^TA).

Aplicaciones de la DVS

Se presentan algunas aplicaciones importantes, luego de recalcar una cuestión. La DVS es terrorífica para cálculos numéricamente estables, ya que U y V son matrices ortogonales. Nunca cambian la longitud de un vector. Debido a que $\|Ux\|^2 = x^TU^TUx = \|x\|^2$, la multiplicación por U no puede destruir el escalamiento.

Por supuesto, Σ puede multiplicar por un gran Σ o bien (lo que es más común) dividir entre un Σ pequeño, y saturar la computadora. Sin embargo, Σ *es la mejor posible*. Revela exactamente qué es grande y qué es pequeño. La razón $\Sigma_{\text{máx}}/\Sigma_{\text{mín}}$ es el **número de condición** de una matriz invertible de n por n. La disponibilidad de esta información es otra razón que explica la conocida popularidad de la DVS. En la segunda aplicación se volverá a esto.

1. Procesamiento de imágenes Suponga que un satélite toma una fotografía, y quiere enviarla a la Tierra. La fotografía puede contener 1000 por 1000 "pixeles": un millón de cuadros pequeños, cada uno con un color definido. Es posible codificar los colores y transmitir 1 000 000 de números. Es mejor encontrar la *información esencial* dentro de la matriz de 1000 por 1000, y enviar sólo eso.

Suponga que se conoce la DVS. La clave está en los valores singulares (en Σ). Típicamente, algunos σs son significantes y otros son extremadamente pequeños. Si se preservan 20 y se desechan 980, entonces sólo se envían las 20 columnas correspondientes de U y V. Las otras 980 columnas se multiplican por $U\Sigma V^T$ por los σs pequeños que se han ignorado. *La multiplicación de matrices puede hacerse como columnas por renglones*:

$$A = U\Sigma V^T = u_1\sigma_1 v_1^T + u_2\sigma_2 v_2^T + \cdots + u_r\sigma_r v_r^T. \quad (3)$$

Cualquier matriz es la suma de r matrices de rango 1. Si sólo se preservan 20 términos, se envían 20 veces 2000 números, en vez de un millón (compresión de 25 a 1).

Las fotografías son verdaderamente sorprendentes, a medida que se incluyen más y más valores. Al principio no se ve nada y repentinamente se reconoce todo. El costo está en el cálculo de la DVS; ésta se ha vuelto mucho más eficaz, aunque es costosa para una matriz grande.

2. El rango efectivo El rango de una matriz es el número de renglones independientes, y el número de columnas independientes. ¡Esto puede ser difícil de decidir en cálculos! En aritmética exacta, el conteo de los pivotes es correcto. La aritmética real puede ser engañosa, aunque eliminar los pivotes pequeños no es la respuesta. Considere lo siguiente:

$$\boldsymbol{\epsilon}\ \textbf{es pequeño} \quad \begin{bmatrix}\epsilon & 2\epsilon\\1 & 2\end{bmatrix} \quad \text{y} \quad \begin{bmatrix}\epsilon & 1\\0 & 0\end{bmatrix} \quad \text{y} \quad \begin{bmatrix}\epsilon & 1\\\epsilon & 1+\epsilon\end{bmatrix}.$$

El rango de la primera es 1, aunque el error por redondeo probablemente produzca un segundo pivote. Ambos pivotes son pequeños, así que ¿cuántos se ignoran? La segunda tiene un pivote pequeño, pero no es posible pretender que este renglón es insignificante. La tercera tiene dos pivotes y su rango es 2, aunque su "rango efectivo" debe ser 1.

A continuación se procede a una medición más estable del rango. El primer paso es utilizar A^TA o AA^T, que son simétricas pero tienen el mismo rango que A. Sus valores característicos –los valores singulares al cuadrado– *no son* engañosos. Con base en la precisión de los datos, se decide en una tolerancia como 10^{-6} y se cuentan los valores singulares arriba de ésta: este es el rango efectivo. Los ejemplos anteriores tienen rango efectivo 1 (cuando ϵ es muy pequeño).

3. Descomposición polar Todo número complejo z es un número positivo r multiplicado por $e^{i\theta}$ en la circunferencia unitaria: $z = re^{i\theta}$. Esto expresa a z en "coordenadas polares". Si se considera que z es una matriz de 1 por 1, r corresponde a una *matriz positiva definida* y $e^{i\theta}$ corresponde a una *matriz ortogonal*. Más exactamente, debido a que $e^{i\theta}$ es complejo y satisface $e^{-i\theta}e^{i\theta} = 1$, forma una *matriz unitaria* de 1 por 1: $U^HU = I$. Se toma el conjugado complejo, así como la traspuesta, de U^H.

La DVS extiende esta "factorización polar" a matrices de cualquier tamaño:

Toda matriz cuadrada real puede factorizarse en $A = QS$, donde Q es **ortogonal** y S es ***semidefinida positiva simétrica***. Si A es invertible, entonces S es positiva definida.

Para demostrar lo anterior simplemente se inserta $V^TV = I$ en medio de la DVS:

$$A = U\Sigma V^T = (UV^T)(V\Sigma V^T). \quad (4)$$

El factor $S = V\Sigma V^T$ es simétrica y semidefinida (porque Σ lo es). El factor $Q = UV^T$ es una matriz ortogonal (porque $Q^TQ = VU^TUV^T = I$). En el caso complejo, S se convierte en hermitiana en vez de simétrica y Q se vuelve unitaria en vez de ortogonal. En el caso invertible, Σ es definida, así como S.

Ejemplo 3 Descomposición polar:

$$A = QS \qquad \begin{bmatrix} 1 & -2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}.$$

Ejemplo 4 Descomposición polar inversa:

$$A = S'Q \qquad \begin{bmatrix} 1 & -2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Los ejercicios muestran cómo, en orden inverso, S cambia pero Q permanece igual. Tanto S como S' son positivas definidas simétricas porque A es invertible.

Aplicación de $A = QS$: Una utilización fundamental de la descomposición polar se encuentra en mecánica del continuo (y recientemente en robótica). En cualquier deformación, es importante separar el alargamiento de la rotación, que es exactamente lo que logra QS. La matriz ortogonal Q es una rotación, y quizá una reflexión. El material no experimenta deformación. La matriz simétrica S tiene valores característicos $\Sigma_1, \ldots, \Sigma_r$, que son los factores de alargamiento (o de compresión). La diagonalización que despliegan estos valores característicos constituye la elección natural de los ejes, denominados ***ejes principales***: así como con las elipses en la sección 6.2. Es S la que requiere trabajar con el material, y almacena energía elástica.

Se observa que S^2 es A^TA, que es positiva definida simétrica cuando A es invertible. S es la raíz cuadrada positiva definida simétrica de A^TA, y Q es AS^{-1}. De hecho, *A puede ser rectangular, en tanto A^TA sea positiva definida.* (Esta es la condición que se ha cumplido, que A tenga columnas independientes.) En el orden inverso $A = S'Q$, la matriz S' es la raíz cuadrada positiva definida simétrica de AA^T.

4. Mínimos cuadrados Para un sistema rectangular $Ax = b$, la solución por mínimos cuadrados proviene de la ecuación normal $A^TA\widehat{x} = A^Tb$. ***Si A tiene columnas dependientes, entonces A^TA no es invertible y $\widehat{x}$ no está determinado.*** Cualquier vector en el espacio nulo puede sumarse a $\widehat{x}$. Ahora es posible completar el capítulo 3, eligiendo un "mejor" (*más corto*) $\widehat{x}$ para toda $Ax = b$.

$Ax = b$ tiene dos posibles dificultades: *renglones dependientes* o *columnas dependientes*. Con renglones dependientes, $Ax = b$ puede no tener solución. Esto ocurre cuando b está fuera del espacio columna de A. En vez de resolver $Ax = b$, se resuelve $A^TA\widehat{x} = A^Tb$. Pero si A tiene columnas dependientes, este $\widehat{x}$ no es único. Es necesario escoger una solución particular de $A^TA\widehat{x} = A^Tb$, y se elige la más corta.

> ***La solución óptima de $Ax = b$ es la solución de menor longitud de $A^TA\widehat{x} = A^Tb$.***

Esa solución de longitud mínima se denomina x^+. Es nuestra opción preferida como la mejor solución de $Ax = b$ (que no tenía solución), y también de $A^TA\widehat{x} = A^Tb$ (que tenía demasiadas). Empezamos con un ejemplo diagonal.

Ejemplo 5 A es diagonal, con renglones dependientes y columnas dependientes:

$$A\widehat{x} = p \quad \text{es} \quad \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \widehat{x}_1 \\ \widehat{x}_2 \\ \widehat{x}_3 \\ \widehat{x}_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}.$$

Todas las columnas terminan con cero. En el espacio columna, el vector más próximo a $b = (b_1, b_2, b_3)$ es $p = (b_1, b_2, 0)$. Lo mejor que puede hacerse con $Ax = b$ es resolver las dos primeras ecuaciones, ya que la tercera ecuación es $0 = b_3$. Este error no puede reducirse, pero los errores en las dos primeras ecuaciones es cero. Así,

$$\widehat{x}_1 = b_1/\sigma_1 \quad \text{y} \quad \widehat{x}_2 = b_2/\sigma_2.$$

Ahora se enfrenta la segunda dificultad. Con la finalidad de hacer a $\widehat{x}$ lo más pequeño posible, se escoge que los totalmente arbitrarios $\widehat{x}_3$ y $\widehat{x}_4$ sean cero. ***La solución de longitud mínima es x^+***:

A^+ es seudoinversa
$x^+ = A^+b$ es más corta

$$x^+ = \begin{bmatrix} b_1/\sigma_1 \\ b_2/\sigma_2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/\sigma_1 & 0 & 0 \\ 0 & 1/\sigma_2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \tag{5}$$

Con esta ecuación se encuentra x^+, y también se despliega *la matriz que produce x^+ a partir de b*. Esta matriz es la ***seudoinversa*** A^+ de la A diagonal. Con base en este ejemplo, Σ^+ y x^+ se conocen para cualquier matriz diagonal Σ:

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \\ & & \end{bmatrix} \qquad \Sigma^+ = \begin{bmatrix} 1/\sigma_1 & & \\ & \ddots & \\ & & 1/\sigma_r \\ & & \end{bmatrix} \qquad \Sigma^+b = \begin{bmatrix} b_1/\sigma_1 \\ \vdots \\ b_r/\sigma_r \\ \end{bmatrix}.$$

La matriz Σ es de m por n, con r elementos diferentes de cero en Σ_i. Su seudoinversa Σ^+ es n por m con elementos r diferentes de cero en $1/\Sigma_i$. **Todos los espacios en blanco son ceros.** Observe que $(\Sigma^+)^+$ es de nuevo Σ. Esto es como $(A^{-1})^{-1} = A$, pero aquí A no es invertible.

A continuación se encuentra x^+ en el caso general. Se afirma que ***la solución más corta x^+, siempre está en el espacio renglón de A.*** Recuerde que cualquier vector $\widehat{x}$ puede separarse en una componente espacio renglón x_r y una componente espacio nulo $\widehat{x} = x_r + x_n$. A continuación se presentan tres cuestiones importantes sobre esa separación:

1. La componente espacio renglón también resuelve $A^TA\widehat{x}_r = A^Tb$, porque $Ax_n = 0$.
2. Las componentes son ortogonales, y cumplen el teorema de Pitágoras:

$$\|\widehat{x}\|^2 = \|x_r\|^2 + \|x_n\|^2, \quad \text{de modo que } \widehat{x} \text{ es más corto cuando } x_n = 0.$$

3. Todas las soluciones de $A^TA\widehat{x} = A^Tb$ tienen el mismo x_r. ***Ese vector es x^+.***

El teorema fundamental del álgebra lineal se mostró en la figura 3.4. Todo p en el espacio columna proviene de uno y sólo un vector x_r en el espacio renglón. *Todo lo que se está haciendo es escoger ese vector $x^+ = x_r$, como la mejor solución de $Ax = b$.*

La seudoinversa en la figura 6.3 empieza con b y regresa a x^+. *Invierte a A donde A es invertible*: entre el espacio renglón y el espacio columna. La seudoinversa elimina al espacio nulo izquierdo al enviarlo a cero, y elimina el espacio nulo escogiendo a x_r como x^+.

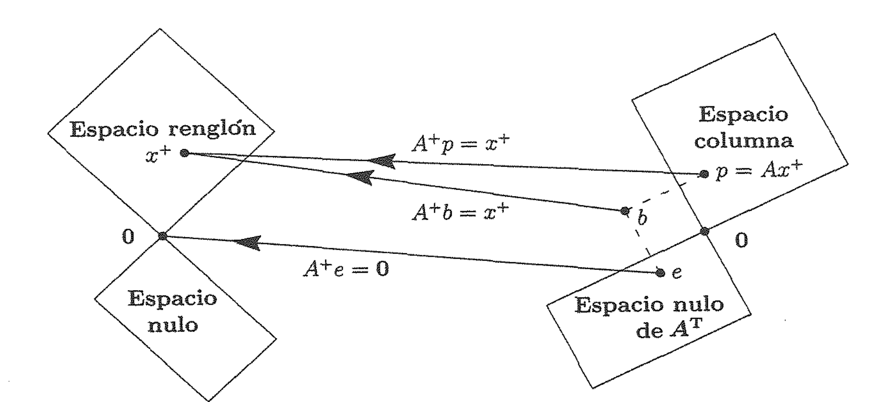


Figura 6.3 La seudoinversa A^+ invierte A donde puede sobre el espacio columna.

Aún no se demuestra que existe una matriz A^+ que siempre proporciona x^+, pero la hay. Esta matriz es de m por n, ya que lleva b y p en $\mathbf{R}^m$ de regreso a x^+ en $\mathbf{R}^n$. Antes de encontrar en general a A^+ se considerará otro ejemplo.

Ejemplo 6 $Ax = b$ es $-x_1 + 2x_2 + 2x_3 = 18$, con un plano completo de soluciones.

Según nuestra teoría, la solución más corta debe estar en el espacio renglón de $A = [-1\ 2\ 2]$. El múltiplo de ese renglón que satisface la ecuación es $x^+ = (-2, 4, 4)$. Hay soluciones más largas como $(-2, 5, 3)$, $(-2, 7, 1)$, o $(-6, 3, 3)$, pero todas tienen componentes diferentes de cero provenientes del espacio nulo. La matriz que produce x^+ a partir de $b = [18]$ es la seudoinversa A^+. Mientras A era de 1 por 3, esta A^+ es de 3 por 1:

$$A^+ = \begin{bmatrix} -1 & 2 & 2 \end{bmatrix}^+ = \begin{bmatrix} -\frac{1}{9} \\ \frac{2}{9} \\ \frac{2}{9} \end{bmatrix} \quad \text{y} \quad A^+[18] = \begin{bmatrix} -2 \\ 4 \\ 4 \end{bmatrix}. \tag{6}$$

El espacio renglón de A es el espacio columna de A^+. A continuación se proporciona una fórmula para A^+:

Si $A = U\Sigma V^T$ (la DVS), entonces su seudoinversa es $A^+ = V\Sigma^+U^T$. (7)

En el ejemplo 6 se tenía $\Sigma = 3$; la raíz cuadrada del valor característico de $AA^T = [9]$. Aquí está de nuevo con Σ y Σ^+:

$$A = \begin{bmatrix} -1 & 2 & 2 \end{bmatrix} = U\Sigma V^T = \begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{bmatrix}$$

$$V\Sigma^+U^T = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} \frac{1}{3} \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{9} \\ \frac{2}{9} \\ \frac{2}{9} \end{bmatrix} = A^+.$$

La solución de longitud mínima por mínimos cuadrados es $x^+ = A^+b = V\Sigma^+U^Tb$.

Demostración la multiplicación por la matriz ortogonal U^T deja sin cambio las longitudes:

$$\|Ax - b\| = \|U\Sigma V^Tx - b\| = \|\Sigma V^Tx - U^Tb\|.$$

Se introduce la nueva incógnita $y = V^Tx = V^{-1}x$, cuya longitud es la misma que x. Así, minimizar $\|Ax - b\|$ es lo mismo que minimizar $\|\Sigma y - U^Tb\|$. Ahora Σ es diagonal y se conoce el mejor y^+. Se trata de $y^+ = \Sigma^+U^Tb$, de modo que el mejor x^+ es Vy^+:

Solución más corta $\qquad x^+ = Vy^+ = V\Sigma^+U^Tb = A^+b.$

Vy^+ está en el espacio renglón, y $A^TAx^+ = A^Tb$ proviene de la **DVS**. ■

Conjunto de problemas 6.3

En los problemas 1 y 2 se calcula la DVS de una matriz singular cuadrada A.

1. Calcule A^TA y sus valores característicos σ_1^2, 0 y sus vectores característicos unitarios v_1, v_2:
$$A = \begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix}.$$

2. *a*) Calcule AA^T y sus valores característicos σ_1^2, 0 y sus vectores característicos unitarios u_1, u_2.
b) Escoja signos de modo que $Av_1 = \Sigma_1 u_1$ y compruebe la DVS:
$$\begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix} = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & 0 \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix}^T.$$
c) ¿Cuáles son los cuatro vectores que proporcionan bases ortonormales para $C(A)$, $N(A)$, $C(A^T)$, $N(A^T)$?

En los problemas 3 a 5 se solicita la DVS de matrices de rango 2.

3. Encuentre la DVS a partir de los vectores característicos v_1, v_2 de A^TA y $Av_i = \sigma_i u_i$:
Matriz de Fibonacci $\qquad A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$

4. Use la parte DVS del demo **eigshow** de MATLAB (o Java en la página del curso `web.mit.edu/18.06`) para encontrar gráficamente los mismos vectores v_1 y v_2.

5. Calcule A^TA y AA^T, así como sus valores característicos y vectores característicos unitarios para
$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$
Multiplique las tres matrices $U\Sigma V^T$ para recuperar A.

Los problemas 6 a 13 abordan las ideas subyacentes de la DVS.

6. Suponga que $u_1, \ldots, u_n$ y $v_1, \ldots, v_n$ son bases ortonormales de $\mathbf{R}^n$. Construya la matriz A que transforma cada v_j en u_j para obtener $Av_1 = u_1, \ldots, Av_n = u_n$.

7. Construya la matriz con rango 1 que tiene $Av = 12u$ para $v = \frac{1}{2}(1, 1, 1, 1)$ y $u = \frac{1}{3}(2, 2, 1)$. Su único valor singular es $\sigma_1 =$ ____.

8. Encuentre $U\Sigma V^T$ si A tiene columnas ortogonales $w_1, \ldots, w_n$ de longitudes $\sigma_1, \ldots, \sigma_n$.

9. Explique cómo $U\Sigma V^T$ expresa A como una suma de r matrices de rango 1 en la ecuación (3): $\qquad A = \sigma_1 u_1 v_1^T + \cdots + \sigma_r u_r v_r^T.$

10. Suponga que A es una matriz simétrica de 2 por 2 con vectores característicos unitarios u_1 y u_2. Si sus valores característicos son $\lambda_1 = 3$ y $\lambda_2 = -2$, ¿cuáles son U, Σ, y V^T?

11. Suponga que A es invertible (con $\sigma_1 > \sigma_2 > 0$). Cambie A por una matriz lo más pequeña posible para obtener una matriz *singular* A_0. Sugerencia: U y V no cambian:

$$\text{Encuentre } A_0\text{, a partir de} \quad A = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix}^T.$$

12. *a*) Si A cambia a $4A$, ¿cuál es el cambio en la DVS?
b) ¿Cuál es la DVS para A^T y A^{-1}?

13. ¿Por qué la DVS para $A + I$ no utiliza simplemente $\Sigma + I$?

14. Encuentre la DVS y la seudoinversa 0^+ de la *matriz cero* de m por n.

15. Encuentre la DVS y la seudoinversa $V\Sigma^+U^T$ de

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \text{y} \quad C = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

16. Si una matriz Q de m por n tiene columnas ortonormales, ¿cuál es Q^+?

17. Diagonalice A^TA para encontrar su raíz cuadrada positiva definida $S = V\Sigma^{1/2}V^T$ y su descomposición polar $A = QS$:

$$A = \frac{1}{\sqrt{10}} \begin{bmatrix} 10 & 6 \\ 0 & 8 \end{bmatrix}.$$

18. ¿Cuál es la solución de longitud mínima por mínimos cuadrados $x^+ = A^+b$ de lo siguiente?

$$Ax = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}.$$

Es posible calcular A^+, o encontrar la solución general para $A^TA\hat{x} = A^Tb$ y escoger la solución que está en el espacio renglón de A. Este problema ajusta el mejor plano $C + Dt + Ez$ a $b = 0$ *y también* a $b = 2$ en $t = z = 0$ (y $b = 2$ en $t = z = 1$).

19. *a*) Si A tiene columnas independientes, su inversa izquierda $(A^TA)^{-1}A^T$ es A^+.
b) Si A tiene renglones independientes, su inversa derecha $A^T(AA^T)^{-1}$ es A^+.
En ambos casos, compruebe que $x^+ = A^+b$ está en el espacio renglón, y que $A^TAx^+ = A^Tb$.

20. Separe $A = U\Sigma V^T$ en su descomposición polar inversa QS'.

21. La expresión $(AB)^+ = B^+A^+$, ¿siempre es verdadera para seudoinversas? Creo que no.

22. Si se eliminan los renglones cero de U se queda con $A = \underline{L}\,\underline{U}$, donde las r columnas de $\underline{L}$ generan el espacio columna de A y los r renglones de $\underline{U}$ generan el espacio renglón. Así, A^+ tiene la fórmula explícita $\underline{U}^T(\underline{U}\,\underline{U}^T)^{-1}(\underline{L}^T\underline{L})^{-1}\underline{L}^T$.

¿Por qué A^+b está en el espacio renglón con $\underline{U}^T$ al frente? ¿Por qué $A^TAA^+b = A^Tb$, de modo que $x^+ = A^+b$ satisface la ecuación normal como debe hacerlo?

23. Explique por qué AA^+ y A^+A son matrices proyección (y por tanto simétricas). ¿Sobre qué subespacios fundamentales se proyectan?

6.4 PRINCIPIOS MINIMALES

En esta sección nos escapamos por primera vez de las ecuaciones lineales. La incógnita x no se proporcionará como la solución de $Ax = b$ o $Ax = \lambda x$. En vez de ello, el vector x se determinará por un principio mínimo.

Es sorprendente la cantidad de leyes naturales que pueden expresarse como principios minimales. Justo el hecho de que los líquidos pesados se hunden hasta el fondo es una consecuencia de minimización de su energía potencial. Y cuando una persona se sienta en un automóvil o se acuesta en una cama, los resortes se ajustan a sí mismos de modo que la energía se minimiza. Una pajilla en un vaso de agua se ve flexionada porque la luz llega al ojo lo más rápido posible. Ciertamente hay ejemplos más intelectuales: el principio fundamental de la ingeniería estructural es la minimización de la energía total.*

Es necesario mencionar de inmediato que estas "energías" no son otra cosa que *funciones cuadráticas positivas definidas*. Y la derivada de una cuadrática es lineal. Volvemos a las conocidas ecuaciones lineales, cuando las primeras derivadas se igualaron a cero. El primer objetivo en esta sección es ***encontrar el principio mínimo equivalente a $Ax = b$, y la minimización equivalente a $Ax = \lambda x$.*** En dimensiones finitas se hará exactamente lo que la teoría de la optimización hace en un problema continuo, donde "primeras derivadas = 0" proporciona una ecuación diferencial. En todo problema, se tiene libertad de resolver la ecuación lineal o de minimizar la cuadrática.

El primer paso es directo: se desea encontrar la "parábola" $P(x)$ cuyo mínimo ocurre cuando $Ax = b$. Si A es justo un escalar, es fácil hacer lo anterior:

$$\text{La gráfica de } \; P(x) = \frac{1}{2}Ax^2 - bx \; \text{ tiene pendiente cero cuando } \; \frac{dP}{dx} = Ax - b = 0.$$

Este punto $x = A^{-1}b$ es un mínimo si A es positiva. Así, la parábola $P(x)$ se abre hacia arriba (véase la figura 6.4). En más dimensiones esta parábola se transforma en un tazón parabólico (un paraboloide). Para asegurar un mínimo de $P(x)$, no un máximo o un punto silla, ¡A debe ser positiva definida!

6H Si A es positiva definida simétrica, entonces $P(x) = \frac{1}{2}x^TAx - x^Tb$ alcanza su mínimo en el punto donde $Ax = b$. En ese punto, $P_{\text{mín}} = -\frac{1}{2}b^TA^{-1}b$.

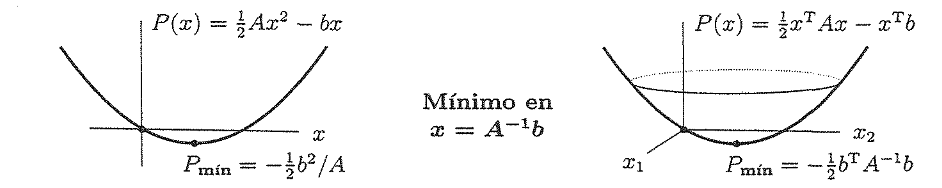


Figura 6.4 La gráfica de una cuadrática positiva $P(x)$ es un tazón parabólico.

*Estoy convencido de que los vegetales y los humanos también se desarrollan según principios minimales. Quizá la civilización está basada en una ley de acción mínima. Debe haber nuevas leyes (y principios minimales) por descubrir en las ciencias sociales y en las ciencias de la vida.

Demostración Suponga que $Ax = b$. Para cualquier vector y, se demostrará que $P(y) \geq P(x)$:

$$\begin{aligned} P(y) - P(x) &= \frac{1}{2}y^{\mathrm{T}}Ay - y^{\mathrm{T}}b - \frac{1}{2}x^{\mathrm{T}}Ax + x^{\mathrm{T}}b \\ &= \frac{1}{2}y^{\mathrm{T}}Ay - y^{\mathrm{T}}Ax + \frac{1}{2}x^{\mathrm{T}}Ax \quad \text{(iguale } b = Ax) \\ &= \frac{1}{2}(y-x)^{\mathrm{T}}A(y-x). \end{aligned} \tag{1}$$

Lo anterior no puede ser positiva definida porque A es positiva definida, y es cero sólo si $y - x = 0$. En todos los demás puntos, $P(y)$ es más grande que $P(x)$, de modo que el mínimo ocurre en x. ■

Ejemplo 1 Minimizar $P(x) = x_1^2 - x_1x_2 + x_2^2 - b_1x_1 - b_2x_2$. El método de costumbre, por cálculo, es igualar a cero las derivadas parciales. Así se obtiene $Ax = b$:

$$\begin{aligned} \partial P/\partial x_1 &= 2x_1 - x_2 - b_1 = 0 \\ \partial P/\partial x_2 &= -x_1 + 2x_2 - b_2 = 0 \end{aligned} \quad \text{significa} \quad \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \tag{2}$$

El álgebra lineal reconoce esta $P(x)$ como $\frac{1}{2}x^{\mathrm{T}}Ax - x^{\mathrm{T}}b$, y sabe de inmediato que $Ax = b$ proporciona el mínimo. Se sustituye $x = A^{-1}b$ en $P(x)$:

$$\textbf{Valor mínimo} \qquad P_{\text{mín}} = \frac{1}{2}(A^{-1}b)^{\mathrm{T}}A(A^{-1}b) - (A^{-1}b)^{\mathrm{T}}b = -\frac{1}{2}b^{\mathrm{T}}A^{-1}b. \tag{3}$$

En aplicaciones, $\frac{1}{2}x^{\mathrm{T}}Ax$ es la energía interna y $-x^{\mathrm{T}}b$ es el trabajo externo. El sistema automáticamente se vuelve $x = A^{-1}b$, donde la energía total $P(x)$ es un mínimo.

Minimización con restricciones

Muchas aplicaciones agregan ecuaciones extra $Cx = d$ por encima del problema de minimización. Estas ecuaciones son **restricciones**. $P(x)$ se minimiza sujeto al requerimiento adicional $Cx = d$. Por lo general x no es capaz de satisfacer n ecuaciones $Ax = b$ y también ℓ restricciones extra $Cx = d$. Se tienen demasiadas ecuaciones y se requieren ℓ incógnitas más.

Estas nuevas incógnitas $y_1, \ldots, y_\ell$ se denominan **multiplicadores de Lagrange**. Integran la restricción en una función $L(x, y)$. Esta fue la brillante aportación de Lagrange:

$$L(x, y) = P(x) + y^{\mathrm{T}}(Cx - d) = \frac{1}{2}x^{\mathrm{T}}Ax - x^{\mathrm{T}}b + x^{\mathrm{T}}C^{\mathrm{T}}y - y^{\mathrm{T}}d.$$

El término en L se escoge exactamente de modo que $\partial L/\partial y = 0$ devuelva $Cx = d$. Cuando las derivadas de L se igualan a cero, se tienen $n + \ell$ ecuaciones para $n + \ell$ incógnitas x y y:

$$\begin{array}{ll} \textbf{Minimización} & \partial L/\partial x = 0: \quad Ax + C^{\mathrm{T}}y = b \\ \textbf{restringida} & \partial L/\partial y = 0: \quad Cx \qquad\quad = d \end{array} \tag{4}$$

Las primeras ecuaciones implican las misteriosas incógnitas y. Bien podría preguntarse qué representan. Estas "incógnitas duales" y indican cuánto el mínimo restringido $P_{C/\text{mín}}$ (que sólo permite a x cuando $Cx = d$) excede al $P_{\text{mín}}$ no restringido (permitiendo todas las x):

$$\textbf{Sensibilidad del mínimo} \qquad P_{C/\text{mín}} = P_{\text{mín}} + \frac{1}{2}y^{\mathrm{T}}(CA^{-1}b - d) \geq P_{\text{mín}}. \tag{5}$$

Ejemplo 2 Suponga que $P(x_1, x_2) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2$. Ciertamente, su valor más pequeño es $P_{\text{mín}} = 0$. Este problema no restringido tiene $n = 2$, $A = I$ y $b = 0$. Así, la ecuación de minimización $Ax = b$ justamente proporciona $x_1 = 0$ y $x_2 = 0$.

Ahora se agrega una restricción $c_1x_1 + c_2x_2 = d$. Esto coloca a x sobre una recta en el plano x_1-x_2. El minimizador anterior $x_1 = x_2 = 0$ no está en la recta. La función lagrangiana $L(x, y) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + y(c_1x_1 + c_2x_2 - d)$ tiene $n + \ell = 2 + 1$ derivadas parciales:

$$\begin{aligned} \partial L/\partial x_1 &= 0 \qquad & x_1 + c_1y &= 0 \\ \partial L/\partial x_2 &= 0 \qquad & x_2 + c_2y &= 0 \\ \partial L/\partial y &= 0 \qquad & c_1x_1 + c_2x_2 &= d. \end{aligned} \tag{6}$$

Al sustituir $x_1 = -c_1y$ y $x_2 = -c_2y$ en la tercera ecuación se obtiene $-c_1^2y - c_2^2y = d$.

$$\textbf{Solución} \qquad y = \frac{-d}{c_1^2 + c_2^2} \qquad x_1 = \frac{c_1d}{c_1^2 + c_2^2} \qquad x_2 = \frac{c_2d}{c_1^2 + c_2^2}. \tag{7}$$

El mínimo restringido de $P = \frac{1}{2}x^{\mathrm{T}}x$ se alcanza en ese punto solución:

$$P_{C/\text{mín}} = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 = \frac{1}{2}\frac{c_1^2d^2 + c_2^2d^2}{(c_1^2 + c_2^2)^2} = \frac{1}{2}\frac{d^2}{c_1^2 + c_2^2}. \tag{8}$$

Lo anterior es igual a $-\frac{1}{2}yd$, como se pronosticaba en la ecuación (5), ya que $b = 0$ y $P_{\text{mín}} = 0$.

En la figura 6.5 se muestra el problema que ha resuelto el álgebra lineal, si las restricciones mantienen a x sobre una recta $2x_1 - x_2 = 5$. Se está buscando **el punto más próximo a (0, 0) sobre esta recta.** La solución es $x = (2, -1)$. Se espera que este vector más corto x sea perpendicular a la recta, y se tiene razón.

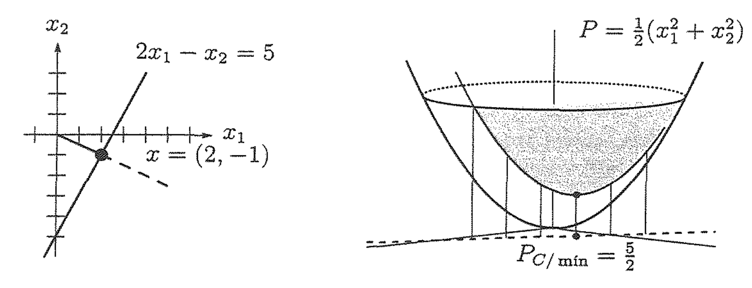


Figura 6.5 Minimización de $\frac{1}{2}\|x\|^2$ para toda x sobre la recta de restricciones $2x_1 - x_2 = 5$.

Mínimos cuadrados de nuevo

En minimización, la gran aplicación son los mínimos cuadrados. El mejor $\widehat{x}$ es el vector que minimiza el error al cuadrado $E^2 = \|Ax - b\|^2$. ¡Esta es una cuadrática y se ajusta a nuestro marco de referencia! Esclarecerá las partes que parecen nuevas:

$$\textbf{Error al cuadrado} \qquad E^2 = (Ax - b)^{\mathrm{T}}(Ax - b) = x^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}x - 2x^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b}. \tag{9}$$

Compare con $\frac{1}{2}x^{\mathrm{T}}Ax - x^{\mathrm{T}}b$ al inicio de esta sección, que condujo a $Ax = b$:

$$[A \text{ cambia a } \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}] \qquad [b \text{ cambia a } A^{\mathrm{T}}\boldsymbol{b}] \qquad [\text{se suma } \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b}].$$

La constante $b^{\mathrm{T}}b$ eleva toda la gráfica, lo cual no afecta al mejor $\widehat{x}$. Los otros dos cambios, A en $A^{\mathrm{T}}A$ y b en $A^{\mathrm{T}}b$, originan una nueva manera de llegar a la ecuación de mínimos cua-

drados (ecuación normal). La ecuación de minimización $Ax = b$ cambia a la

Ecuación por mínimos cuadrados $$A^{\mathrm{T}}A\widehat{x} = A^{\mathrm{T}}b. \quad (10)$$

El tema de la optimización requiere todo un libro. Nos detendremos mientras sea álgebra lineal pura.

El cociente de Rayleigh

El segundo objetivo es encontrar un problema de minimización que sea equivalente a $Ax = \lambda x$, lo cual no es tan fácil. La función a minimizar no puede ser cuadrática, ya que entonces su derivada tendría que ser lineal, y el problema de valores característicos es no lineal (λ multiplicado por x). El truco exitoso es dividir una cuadrática entre otra:

Cociente de Rayleigh Minimizar $$R(x) = \frac{x^{\mathrm{T}}Ax}{x^{\mathrm{T}}x}.$$

6I Principio de Rayleigh: El valor mínimo del cociente de Rayleigh es el menor valor característico λ_1. $R(x)$ alcanza el mínimo en el primer vector característico x_1 de A:

Mínimo donde $Ax_1 = \lambda x_1$ $$R(x_1) = \frac{x_1^{\mathrm{T}}Ax_1}{x_1^{\mathrm{T}}x_1} = \frac{x_1^{\mathrm{T}}\lambda_1 x_1}{x_1^{\mathrm{T}}x_1} = \lambda_1.$$

Si se mantiene $x^{\mathrm{T}}Ax = 1$, entonces $R(x)$ es un mínimo cuando $x^{\mathrm{T}}x = \|x\|^2$ es lo más grande posible. Se está buscando el punto sobre el elipsoide $x^{\mathrm{T}}Ax = 1$ lo más lejano posible al origen: el vector x de mayor longitud. Con base en el análisis que se hizo sobre el elipsoide, su eje mayor apunta a lo largo del primer vector característico. Así, $R(x)$ es un mínimo en x_1.

Algebraicamente es posible diagonalizar la matriz simétrica A por medio de una matriz ortogonal: $Q^{\mathrm{T}}AQ = \Lambda$. Luego se hace $x = Qy$ y el cociente se facilita:

$$R(x) = \frac{(Qy)^{\mathrm{T}}A(Qy)}{(Qy)^{\mathrm{T}}(Qy)} = \frac{y^{\mathrm{T}}\Lambda y}{y^{\mathrm{T}}y} = \frac{\lambda_1 y_1^2 + \cdots + \lambda_n y_n^2}{y_1^2 + \cdots + y_n^2}. \quad (11)$$

El mínimo de R es λ_1, en el punto en que $y_1 = 1$ y $y_2 = \cdots = y_n = 0$:

En todos los puntos $$\lambda_1(y_1^2 + y_2^2 + \cdots + y_n^2) \leq (\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2).$$

El cociente de Rayleigh en la ecuación (11) ***nunca es menor que λ_1*** y ***nunca es mayor que λ_n*** (el mayor valor característico). Su mínimo está en el vector característico x_1 y su máximo está en x_n:

Máximo donde $Ax_n = \lambda_n x_n$ $$R(x_n) = \frac{x_n^{\mathrm{T}}Ax_n}{x_n^{\mathrm{T}}x_n} = \frac{x_n^{\mathrm{T}}\lambda_n x_n}{x_n^{\mathrm{T}}x_n} = \lambda_n.$$

Un detalle pequeño pero importante: el cociente de Rayleigh es igual a a_{11}, cuando el vector ensayo es $x = (1, 0, \ldots, 0)$. Así, a_{11} (sobre la diagonal principal) está entre λ_1 y λ_n. Esto puede verse en la figura 6.6, donde la distancia horizontal a la elipse (donde $a_{11}x^2 = 1$) está entre la distancia más corta y la distancia más larga:

$$\frac{1}{\sqrt{\lambda_n}} \leq \frac{1}{\sqrt{a_{11}}} \leq \frac{1}{\sqrt{\lambda_1}} \quad \text{que es} \quad \lambda_1 \leq a_{11} \leq \lambda_n.$$

Los elementos diagonales de cualquier matriz simétrica están entre λ_1 y λ_n. Para ver más claramente lo anterior, la figura 6.6 se trazó para una matriz positiva definida de 2 por 2.

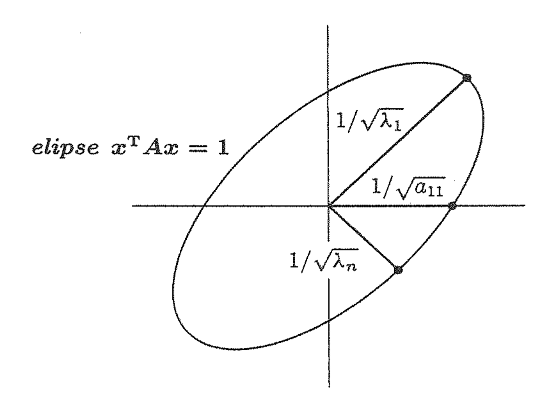


Figura 6.6 Ambas, la $x = x_1/\sqrt{\lambda_1}$ más alejada y la $x = x_n/\sqrt{\lambda_n}$ más próxima proporcionan $x^{\mathrm{T}}Ax = x^{\mathrm{T}}\lambda x = 1$. Estos son los ejes mayor y menor de la elipse.

Entrelazamiento de los valores característicos

Los vectores característicos intermedios $x_2, \ldots, x_{n-1}$ son ***puntos silla*** del cociente de Rayleigh (derivadas cero, aunque ningún mínimo o máximo). La dificultad con estos puntos es que no se tiene ni idea de si $R(x)$ está arriba o abajo de ellos. Esto hace más difícil de estimar a los valores característicos intermedios $\lambda_2, \ldots, \lambda_{n-1}$.

Para este tópico opcional, la clave consiste en encontrar un mínimo o un máximo restringido. Las restricciones provienen de la propiedad básica de las matrices simétricas: x_j es perpendicular a los otros vectores característicos.

6J El mínimo de $R(x)$ sujeto a $x^{\mathrm{T}}x_1 = 0$ es λ_2. El mínimo de $R(x)$ sujeto a cualquiera otra restricción $x^{\mathrm{T}}v = 0$ no está por arriba de λ_2:

$$\lambda_2 = \min_{x^{\mathrm{T}}x_1=0} R(x) \quad \text{y} \quad \lambda_2 \geq \min_{x^{\mathrm{T}}v=0} R(x). \quad (12)$$

Este "principio máximo" hace de λ_2 el *máximo sobre todas las v del mínimo* de $R(x)$ con $x^{\mathrm{T}}v = 0$. Esto supone un método para estimar λ_2 sin conocer λ_1.

Ejemplo 3 Eliminar el último renglón y la última columna de cualquier matriz simétrica:

$$\begin{matrix}\lambda_1(A) = 2-\sqrt{2}\\ \lambda_2(A) = 2\\ \lambda_3(A) = 2+\sqrt{2}\end{matrix} \quad A = \begin{bmatrix} 2 & -1 & 0\\ -1 & 2 & -1\\ 0 & -1 & 2\end{bmatrix} \text{ se convierte en } B = \begin{bmatrix} 2 & -1\\ -1 & 2\end{bmatrix} \quad \begin{matrix}\lambda_1(B) = 1\\ \lambda_2(B) = 3.\end{matrix}$$

El segundo valor característico $\lambda_2(A) = 2$ está por arriba del menor valor característico $\lambda_1(B) = 1$. El menor valor característico $\lambda_1(A) = 2 - \sqrt{2}$ *está por abajo de* $\lambda_1(B)$. Así, $\lambda_1(B)$ está atrapado entre ambos.

En este ejemplo se escogió $v = (0, 0, 1)$, de modo que la restricción $x^{\mathrm{T}}v = 0$ eliminó la tercera componente de x (reduciendo así A a B).

La fotografía completa es un ***entrelazamiento de valores característicos***:

$$\lambda_1(A) \le \lambda_1(B) \le \lambda_2(A) \le \lambda_2(B) \le \cdots \le \lambda_{n-1}(B) \le \lambda_n(A). \qquad (13)$$

Esto tiene una interpretación natural para un elipsoide, cuando es cortado por un plano que pasa por el origen. La sección transversal es un elipsoide de una dimensión menor. El eje mayor de esta sección transversal no puede ser más largo que el eje mayor de todo el elipsoide: $\lambda_1(B) \ge \lambda_1(A)$. Sin embargo, el eje mayor de la sección transversal es *por lo menos tan largo como el segundo eje* del elipsoide original: $\lambda_1(B) \le \lambda_2(A)$. De manera semejante, el eje menor de la sección transversal es menor que el segundo eje original, y mayor que el eje menor original: $\lambda_2(A) \le \lambda_2(B) \le \lambda_3(A)$.

Lo mismo puede verse en mecánica. Cuando los resortes y las masas están oscilando, suponga que una masa se mantiene en equilibrio. Entonces la menor frecuencia se incrementa, pero no por arriba de λ_2. La mayor frecuencia disminuye, pero no por debajo λ_{n-1}.

Se termina con tres observaciones. Espero que su intuición le indique que son correctas.

Observación 1 El **principio máximo** se extiende a subespacios S_j de dimensión j:

$$\textbf{Máximo del mínimo} \qquad \lambda_{j+1} = \max_{\text{todas las } S_j} \left[\min_{x \perp S_j} R(x) \right]. \qquad (14)$$

Observación 2 Para λ_{n-j} también hay un **principio minimax**:

$$\textbf{Mínimo del máximo} \qquad \lambda_{n-j} = \min_{\text{todas las } S_j} \left[\max_{x \perp S_j} R(x) \right]. \qquad (15)$$

Si $j = 1$, se está maximizando $R(x)$ sobre una restricción $x^Tv = 0$. Este máximo está entre λ_{n-1} y λ_n no restringidas. La restricción más fuerte hace que x sea perpendicular al vector característico superior $v = x_n$. Así, el mejor x es el siguiente vector característico x_{n-1}. El "mínimo de los máximos" es λ_{n-1}.

Observación 3 Para el problema generalizado $Ax = \lambda Mx$, los mismos principios se cumplen si M es positiva definida. En el cociente de Rayleigh, x^Tx se convierte en x^TMx:

$$\textbf{Cociente de Rayleigh} \quad \text{Al minimizar } R(x) = \frac{x^TAx}{x^TMx} \text{ se obtiene } \lambda_1(M^{-1}A). \qquad (16)$$

Incluso para masas *desiguales* en un sistema oscilatorio ($M \ne I$), al mantener una masa en equilibrio se eleva la menor frecuencia y se disminuye la mayor frecuencia.

Conjunto de problemas 6.4

1. Considere el sistema $Ax = b$ dado por

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 4 \end{bmatrix}.$$

Construya la cuadrática correspondiente $P(x_1, x_2, x_3)$, calcule sus derivadas parciales $\partial P/\partial x_i$, y compruebe que se hacen cero exactamente en la solución deseada.

2. Complete al cuadrado en $P = \frac{1}{2}x^TAx - x^Tb = \frac{1}{2}(x - A^{-1}b)^TA(x - A^{-1}b)$ + constante. Esto es igual a $P_{\text{mín}}$ porque el término anterior nunca es negativo. (¿Por qué?)

3. Encuentre el mínimo, en caso de haberlo, de $P_1 = \frac{1}{2}x^2 + xy + y^2 - 3y$ y $P_2 = \frac{1}{2}x^2 - 3y$. ¿Qué matriz A está asociada con P_2?

4. (Repaso) Otra cuadrática que ciertamente tiene su mínimo en $Ax = b$ es

$$Q(x) = \frac{1}{2}\|Ax - b\|^2 = \frac{1}{2}x^TA^TAx - x^TA^Tb + \frac{1}{2}b^Tb.$$

Al comparar Q con P, e ignorar la constante $\frac{1}{2}b^Tb$, ¿qué sistema de ecuaciones se obtiene en el mínimo de Q? ¿Cómo se denominan estas ecuaciones en la teoría de mínimos cuadrados?

5. Para cualquier matriz simétrica A, calcule la razón $R(x)$ para la elección especial $x = (1, \ldots, 1)$. ¿Cómo está relacionada la suma de todos los elementos a_{ij} con λ_1 y λ_n?

6. Con $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$, encuentre una elección de x con la que se obtenga un menor $R(x)$ que la cota $\lambda_1 \le 2$ que proviene de los elementos diagonales. ¿Cuál es el valor mínimo de $R(x)$?

7. Si B es positiva definida, a partir del cociente de Rayleigh demuestre que el menor valor característico de $A + B$ es mayor que el menor valor característico de A.

8. Si λ_1 y μ_1 son los menores valores característicos de A y B, demuestre que el menor valor característico θ_1 de $A + B$ es por lo menos tan grande como $\lambda_1 + \mu_1$. (Intente el correspondiente vector característico x en los cocientes de Rayleigh.)

Nota Quizá los problemas 7 y 8 son los resultados más típicos y más importantes que resultan fácilmente del principio de Rayleigh, pero no es así en el caso de las ecuaciones de valores característicos.

9. Si B es positiva definida, a partir del principio minimax (12) demuestre que el *segundo* menor valor característico se incrementa al sumar B : $\lambda_2(A + B) > \lambda_2(A)$.

10. Si se eliminan *dos* renglones y *dos* columnas de A, ¿qué desigualdades son de esperar entre el menor valor característico μ de la nueva matriz y los λs originales?

11. Encuentre los valores mínimos de

$$R(x) = \frac{x_1^2 - x_1x_2 + x_2^2}{x_1^2 + x_2^2} \quad \text{y} \quad R(x) = \frac{x_1^2 - x_1x_2 + x_2^2}{2x_1^2 + x_2^2}.$$

12. A partir de la ecuación (11), demuestre que $R(x)$ nunca es mayor que el mayor valor característico λ_n.

13. El principio minimax para λ_j implica subespacios S_j de dimensión j:

$$\textbf{Equivalente a la ecuación (15)} \qquad \lambda_j = \min_{S_j} \left[\max_{x \text{ en } S_j} R(x) \right].$$

a) Si λ_j es positivo, deduzca que todo S_j contiene un vector x con $R(x) > 0$.
b) Deduzca que S_j contiene un vector $y = C^{-1}x$ con $y^TC^TACy/y^Ty > 0$.
c) Concluya que el j-ésimo valor característico de C^TAC, con base en el principio minimax, también es positivo, suponiendo de nuevo la *ley de la inercia* proporcionada en la sección 6.2.

14. Demuestre que el menor valor característico λ_1 de $Ax = \lambda Mx$ no es más largo que la razón a_{11}/m_{11} de los elementos en las esquinas.

15. ¿Qué subespacio particular S_2 en el problema 13 proporciona el valor mínimo λ_2? En otras palabras, ¿sobre qué S_2 se cumple que el máximo de $R(x)$ es igual a λ_2?

16. (Recomendado) A partir de la submatriz cero, decida los signos de los n valores característicos:

$$A = \begin{bmatrix} 0 & \cdot & 0 & 1 \\ \cdot & \cdot & 0 & 2 \\ 0 & 0 & 0 & \cdot \\ 1 & 2 & \cdot & n \end{bmatrix}.$$

17. (Mínimo restringido) Suponga que el mínimo sin restringir $x = A^{-1}b$ satisface la restricción $Cx = d$. Compruebe que la ecuación (5) proporciona correctamente $P_{C/\text{mín}} = P_{\text{mín}}$; el término de corrección es cero.

6.5 EL MÉTODO DEL ELEMENTO FINITO

En la sección previa sobre principios mínimos se presentaron dos conceptos fundamentales:

i) Resolver $Ax = b$ es equivalente a minimizar $P(x) = \frac{1}{2}x^TAx - x^Tb$.

ii) Resolver $Ax = \lambda_1 x$ es equivalente a minimizar $R(x) = x^TAx/x^Tx$.

A continuación se intentará explicar cómo es posible aplicar estas ideas.

La historia es larga, ya que estos principios se conocen desde hace más de un siglo. En ingeniería, problemas como el doblado de planchas, o problemas de física como el estado fundamental (función característica) de un átomo, la minimización se ha utilizado para obtener una aproximación tosca a la solución verdadera. La aproximación *tenía* que ser tosca; las computadoras eran humanas. Los principios i) y ii) ahí estaban, aunque no era posible implementarlos.

Resulta evidente que la computadora conllevaría una revolución. Lo que saltó hacia delante fue el método de diferencias finitas, ya que es fácil "discretizar" una ecuación diferencial. Ya en la sección 1.7, las derivadas se sustituyeron por diferencias. La región física es cubierta con una malla, y $u'' = f(x)$ se convirtió en $u_{j+1} - 2u_j + u_{j-1} = h^2 f_j$. En la década de 1950 surgieron nuevos métodos para resolver sistemas $Au = f$ que eran muy grandes y ralos: actualmente los algoritmos y el hardware son mucho más rápidos.

Lo que no se reconoció por completo fue que incluso las diferencias finitas se vuelven increíblemente complicadas para problemas ingenieriles reales, como las deformaciones que experimenta un avión. *La verdadera dificultad no estriba en resolver las ecuaciones, sino en plantearlas.* Para una región irregular el trozo de malla se colocó junto con triángulos, cuadriláteros o tetraedros. Luego se requirió un método sistemático para aproximar las leyes físicas subyacentes. La computadora tuvo que ayudar no sólo en la solución de $Au = f$ y $Ax = \lambda x$, sino también en su planteamiento.

Puede conjeturarse lo que ocurrió. Los antiguos métodos están de vuelta, con un nuevo concepto y un nuevo nombre. El nuevo nombre es ***método del elemento finito***. La nueva idea utiliza más de la potencia de la computadora —en la construcción de una aproximación discreta, en su resolución y en el despliegue de los resultados— que cualquiera otra técnica en la computación científica.* Si la idea básica es simple, las aplicaciones pueden ser complicadas.

*Por favor, disculpe este entusiasmo; sé que el método puede no ser eterno.

Para problemas en esta escala, el único punto que no puede discutirse es su costo; me temo que mil millones de dólares es una estimación conservadora del costo hasta la fecha. Espero que algunos lectores tengan el vigor necesario para dominar el método del elemento finito y utilizarlo correctamente.

Funciones ensayo

Empezando con el clásico ***principio de Rayleigh-Ritz***, se introducirá la nueva idea de los elementos finitos. La ecuación puede ser $-u'' = f(x)$ con condiciones en la frontera $u(0) = u(1) = 0$. Este problema es de *dimensión infinita* (el vector b se sustituyó por una función f, y la matriz A se convierte en $-d^2/dx^2$). Es posible escribir la energía cuyo mínimo se busca, al sustituir los productos internos v^Tf por integrales de $v(x)f(x)$:

$$\textbf{Energía total} \quad P(v) = \frac{1}{2}v^TAv - v^Tf = \frac{1}{2}\int_0^1 v(x)(-v''(x))\,dx - \int_0^1 v(x)f(x)\,dx. \quad (1)$$

$P(v)$ debe minimizarse sobre todas las funciones $v(x)$ que satisfacen $v(0) = v(1) = 0$. ***La función que proporciona el mínimo es la solución $u(x)$.*** La ecuación diferencial se ha convertido en un principio mínimo, y sólo queda integrar por partes:

$$\int_0^1 v(-v'')\,dx = \int_0^1 (v')^2\,dx - [vv']_{x=0}^{x=1} \text{ de modo que } P(v) = \int_0^1 \left[\frac{1}{2}(v'(x))^2 + v(x)f(x)\right]dx.$$

El término vv' es cero en ambos límites, ya que v lo es. Así, $\int (v'(x))^2\,dx$ es *positiva*, como x^TAx. Se tiene garantizado un mínimo.

El cálculo exacto del mínimo es equivalente a resolver exactamente la ecuación diferencial. *El principio de Rayleigh-Ritz produce un problema n-dimensional al escoger sólo n funciones ensayo* $V_1(x), \ldots, V_n(x)$. A partir de todas las combinaciones $V = y_1V_1(x) + \cdots + y_nV_n(x)$ se busca la combinación particular (se denomina U) que minimiza $P(V)$. Esta es la idea clave: minimizar sobre un subespacio de Vs en vez de sobre todas las $v(x)$ posibles. La función que proporciona el mínimo es $U(x)$. Es deseable y de esperar que $U(x)$ esté próxima a la $u(x)$ correcta.

Al sustituir V por v, la cuadrática se convierte en

$$P(V) = \frac{1}{2}\int_0^1 (y_1V_1'(x) + \cdots + y_nV_n'(x))^2\,dx - \int_0^1 (y_1V_1(x) + \cdots + y_nV_n(x))f(x)\,dx. \quad (2)$$

Las funciones ensayo V se escogen de antemano. ¡Este es el paso clave! Las incógnitas $y_1, \ldots, y_n$ van en un vector y. Luego, $P(V) = \frac{1}{2}y^TAy - y^Tb$ se reconoce como una de las cuadráticas de costumbre. Los elementos A_{ij} de la matriz son $\int V_i'V_j'\,dx$ = coeficientes de y_iy_j. Las componentes b_j son $\int V_jf\,dx$. Ciertamente, es posible encontrar el mínimo de $\frac{1}{2}y^TAy - y^Tb$ al resolver $Ay = b$. En consecuencia, el método de Rayleigh-Ritz consta de tres pasos:

1. *Escoger las funciones ensayo* $V_1, \ldots, V_n$.
2. *Calcular los coeficientes* A_{ij} *y* b_j.
3. *Resolver* $Ay = b$ para encontrar $U(x) = y_1V_1(x) + \cdots + y_nV_n(x)$.

Todo depende del paso 1. A menos que las funciones $V_j(x)$ sean extremadamente simples, los otros pasos serían virtualmente imposibles. Y a menos que alguna combinación de los V_j esté próxima a la solución $u(x)$ verdadera, estos pasos serán inútiles. Para combinar la facilidad de hacer los cálculos y la exactitud, ***la idea clave que hace exitosos los elementos finitos, es el uso de los polinomios por partes como las funciones ensayo $V(x)$.***

Elementos finitos lineales

El elemento finito más simple y de mayor uso es **lineal por partes**. En los puntos interiores $x_1 = h, x_2 = 2h, \ldots, x_n = nh$ se colocan nodos, así como para las diferencias finitas. Luego, V_j es la "función sombrero" que es igual a 1 en el nodo x_j, y cero en todos los demás nodos (véase la figura 6.7*a*). Está concentrado en un pequeño intervalo alrededor de su nodo, y es cero en todas las demás partes (incluyendo $x = 0$ y $x = 1$). Cualquier combinación $y_1V_1 + \cdots + y_nV_n$ debe tener el valor y_j en el nodo j (los otros Vs son cero ahí), de modo que resulta fácil trazar su gráfica (véase la figura 6.7*b*).

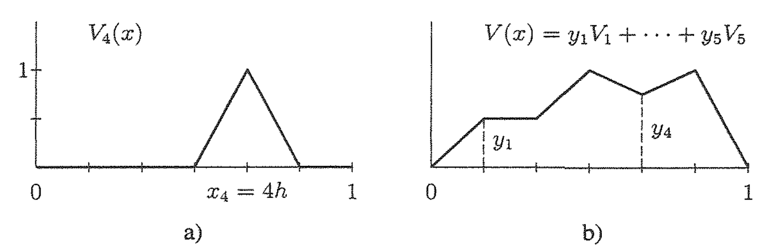


Figura 6.7 Funciones sombrero y sus combinaciones lineales.

En el paso 2 se calculan los coeficientes $A_{ij} = \int V_i'V_j'\,dx$ en la "matriz de rigidez" A. La pendiente V_j' es igual a $1/h$ en el pequeño intervalo a la izquierda de x_j, y a $-1/h$ en el intervalo a la derecha. *Si estos "intervalos dobles" no se traslapan, el producto $V_i'V_j'$ es cero y $A_{ij} = 0$.* Cada función sombrero se traslapa consigo misma y con sólo dos vecinos:

En la diagonal $i = j$ $$A_{ii} = \int V_i'V_i'\,dx = \int \left(\frac{1}{h}\right)^2 dx + \int \left(-\frac{1}{h}\right)^2 dx = \frac{2}{h}.$$

Fuera de la diagonal $i = j \pm 1$ $$A_{ij} = \int V_i'V_j'\,dx = \int \left(\frac{1}{h}\right)\left(\frac{-1}{h}\right) dx = \frac{-1}{h}.$$

Así, la matriz de rigidez es realmente tridiagonal:

$$\textbf{Matriz de rigidez} \quad A = \frac{1}{h}\begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix}.$$

¡Esto se ve como diferencias finitas! Ha conducido a miles de discusiones sobre la relación entre estos dos métodos. Elementos finitos más complicados —polinomios de grado superior, definidos sobre triángulos o cuadriláteros para ecuaciones diferenciales parciales— también producen matrices ralas. Podría pensarse que los elementos finitos es una forma sistemática para construir ecuaciones en diferencias exactas sobre mallas irregulares. La cuestión esencial es la *sencillez* de estos polinomios por partes. Dentro de cada elemento, sus pendientes son fáciles de calcular e integrar.

Las componentes b_j en el miembro derecho son nuevas. En vez de simplemente contar con el valor de f en x_j, como para las diferencias finitas, ahora son un promedio de f alrededor de ese punto: $b_j = \int V_j f\,dx$. Luego, en el paso 3, se resuelve el sistema tridiagonal $Ay = b$, que proporciona los coeficientes en la función ensayo minimizadora $U = y_1V_1 + \cdots + y_nV_n$. Al unir todas estas alturas y_j con una recta quebrada, se obtiene la solución aproximada $U(x)$.

Ejemplo 1 $-u'' = 2$ con $u(0) = u(1) = 0$, y solución $u(x) = x - x^2$.

La aproximación usa tres intervalos y dos funciones sombrero, con $h = \frac{1}{3}$. La matriz A es de 2 por 2. El miembro derecho requiere integración de la función sombrero multiplicada por $f(x) = 2$. Lo anterior produce el doble del área $\frac{1}{3}$ bajo el sombrero:

$$A = 3\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \quad \text{y} \quad b = \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}.$$

La solución de $Ay = b$ es $y = (\frac{2}{9}, \frac{2}{9})$. La mejor $U(x)$ es $\frac{2}{9}V_1 + \frac{2}{9}V_2$, que es igual a $\frac{2}{9}$ en los puntos de la malla. *Esto coincide con la solución exacta* $u(x) = x - x^2 = \frac{1}{3} - \frac{1}{9}$.

En un ejemplo más complicado, la aproximación no es exacta en los nodos. Aunque está extraordinariamente próxima. La teoría subyacente es explicada en el libro del autor *An Analysis of the Finite Element Method* (consulte la página `www.wellesleycambridge.com`) escrito al alimón con George Fix. En otros libros se proporcionan aplicaciones más detalladas, y el tema de los elementos finitos se ha convertido en una parte importante de la educación de la ingeniería. Se aborda en *Introduction to Applied Mathematics*, y también en mi nuevo libro *Applied Mathematics and Scientific Computing*. Ahí se analizan ecuaciones diferenciales parciales, donde el método realmente se encuentra en su medio.

Problemas de valores característicos

El concepto de Rayleigh-Ritz –minimizar sobre una familia de dimensión finita de Vs en lugar de hacerlo sobre todas las vs admisibles– también es de utilidad para los problemas de valores característicos. El verdadero mínimo del cociente de Rayleigh es la frecuencia fundamental λ_1. Su mínimo aproximado Λ_1 es más grande, ya que la clase de funciones ensayo se ha restringido a los Vs. Este paso era completamente natural e inevitable: aplicar las nuevas ideas del elemento finito a esta forma variacional del problema de valores característicos establecida hace bastante.

El mejor ejemplo de un problema de valores característicos tiene $u(x) = \operatorname{sen} \pi x$ y $\lambda_1 = \pi^2$:

Función característica $u(x)$ $\quad -u'' = \lambda u, \quad$ con $\quad u(0) = u(1) = 0.$

Esta función sen πx minimiza el cociente de Rayleigh $v^{\mathrm{T}}Av/v^{\mathrm{T}}v$:

$$\textbf{Cociente de Rayleigh} \quad R(v) = \frac{\int_0^1 v(x)(-v''(x))\,dx}{\int_0^1 (v(x))^2\,dx} = \frac{\int_0^1 (v'(x))^2\,dx}{\int_0^1 (v(x))^2\,dx}.$$

Esta es una razón de energía potencial a energía cinética, y están en equilibrio en el vector característico. Normalmente este vector característico es desconocido, y para aproximarlo es necesario admitir sólo los candidatos de ensayo $V = y_1V_1 + \cdots + y_nV_n$:

$$R(V) = \frac{\int_0^1 (y_1V_1' + \cdots + y_nV_n')^2\,dx}{\int_0^1 (y_1V_1 + \cdots + y_nV_n)^2\,dx} = \frac{y^{\mathrm{T}}Ay}{y^{\mathrm{T}}My}.$$

A continuación se aborda el problema matricial: Minimizar $y^{\mathrm{T}}Ay/y^{\mathrm{T}}My$. Con $M = I$, lo anterior conduce al problema de valores característicos estándar $Ay = \lambda y$. Sin embargo, la matriz M es tridiagonal, ya que la aproximación a esta función se traslapa. Es exactamente la situación que conlleva el *problema generalizado de valores característicos*. **El valor mínimo Λ_1 es el menor valor característico de $Ay = \lambda My$.** Esta Λ_1 está próxima a (y por arriba de) π^2. El vector característico y proporciona la aproximación $U = y_1V_1 + \cdots + y_nV_n$ a la función característica.

Así como en el problema de estática, el método puede resumirse en tres pasos: 1) escoger los V_j, 2) calcular A y M, y 3) resolver $Ay = \lambda My$. No sé por qué esto cuesta mil millones de dólares.

Conjunto de problemas 6.5

1. Use tres funciones sombrero, con $h = \frac{1}{4}$, para resolver $-u'' = 2$ con $u(0) = u(1) = 0$. Compruebe que la aproximación U coincide con $u = x - x^2$ en los nodos.

2. Resuelva $-u'' = x$ con $u(0) = u(1) = 0$. Luego resuelva aproximadamente con dos funciones sombrero y $h = \frac{1}{3}$. ¿Dónde está el error más grande?

3. Suponga que $-u'' = 2$, con la condición en la frontera $u(1) = 0$ cambiada a $u'(1) = 0$. Esta condición "natural" sobre u' no requiere ser impuesta sobre las funciones ensayo V. Con $h = \frac{1}{3}$, existe un *semisombrero* extra V_3, que va de 0 a 1 entre $x = \frac{2}{3}$ y $x = 1$. Calcule $A_{33} = \int (V_3')^2\,dx$ y $f_3 = \int 2V_3\,dx$. Resuelva $Ay = f$ para la solución del elemento finito $y_1V_1 + y_2V_2 + y_3V_3$.

4. Resuelva $-u'' = 2$ con una sola función sombrero, pero coloque su nodo en $x = \frac{1}{4}$ en vez de $x = \frac{1}{2}$. (Trace esta función V_1.) Con condiciones a la frontera $u(0) = u(1) = 0$, compare la aproximación por elementos finitos con la $u = x - x^2$ verdadera.

5. El *método de Galerkin* empieza con la ecuación diferencial (por ejemplo $-u'' = f(x)$) en vez de con la energía P. La solución ensayo sigue siendo $u = y_1V_1 + y_2V_2 + \cdots + y_nV_n$, y las ys se escogen para hacer que la diferencia entre $-u''$ y f sea ortogonal para todo V_j:

$$\textbf{Galerkin} \qquad \int (-y_1V_1'' - y_2V_2'' - \cdots - y_nV_n'')V_j\,dx = \int f(x)V_j(x)\,dx.$$

Integre por partes el miembro izquierdo para llegar a $Ay = f$, demostrando que *el método de Galerkin proporciona las mismas A y f que Rayleigh–Ritz para problemas simétricos.*

6. Una identidad fundamental para cuadráticas demuestra $y = A^{-1}b$ como la minimización de:

$$P(y) = \frac{1}{2}y^TAy - y^Tb = \frac{1}{2}(y - A^{-1}b)^TA(y - A^{-1}b) - \frac{1}{2}b^TA^{-1}b.$$

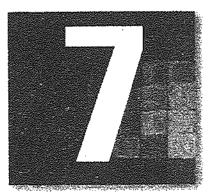
El mínimo sobre un *subespacio* de funciones ensayo está en la *y más próxima* a $A^{-1}b$. (Esto hace lo más pequeño posible al término de la derecha; constituye la clave para la convergencia de U a u.) Si $A = I$ y $b = (1, 0, 0)$, ¿qué múltiplo de $V = (1, 1, 1)$ proporciona el menor valor para $P(y) = \frac{1}{2}y^Ty - y_1$?

7. Para una simple función sombrero $V(x)$ centrada en $x = \frac{1}{2}$, calcule $A = \int (V')^2\,dx$ y $M = \int V^2\,dx$. En el problema de valores característicos de 1 por 1, ¿$\lambda = A/M$ es mayor o menor que el verdadero valor característico $\lambda = \pi^2$?

8. Para las funciones V_1 y V_2 centradas en $x = h = \frac{1}{3}$ y $x = 2h = \frac{2}{3}$, calcule la matriz masa de 2 por 2 $M_{ij} = \int V_iV_j\,dx$, y resuelva el problema de valores característicos $Ax = \lambda Mx$.

9. ¿Cuál es la matriz masa $M_{ij} = \int V_iV_j\,dx$ para n funciones sombrero con $h = \frac{1}{n+1}$?

Capítulo



Cálculos con matrices

7.1 INTRODUCCIÓN

Uno de los objetivos de este libro es explicar las partes útiles de la teoría de matrices. En comparación con libros más antiguos sobre álgebra lineal abstracta, la teoría subyacente no ha cambiado radicalmente. Una de las mejores cosas sobre el tema es que la teoría es realmente esencial para las aplicaciones. Lo que es diferente es el *cambio en el énfasis* que viene con un nuevo punto de vista. La eliminación se convierte en más que justo una forma de encontrar una base para el espacio renglón, y el proceso de Gram-Schmidt no sólo es una demostración de que todo subespacio tiene una base ortonormal. En vez de lo anterior, realmente se *necesitan* estos algoritmos. Y se requiere una descripción conveniente, $A = LU$ o $A = QR$, de lo que hacen aquéllos.

Este capítulo avanzará unos pasos más en la misma dirección. Supongo que estos pasos están regidos por la necesidad computacional, más que por la elegancia, y no sé dónde pedir disculpas por esto; los hace sentir muy superficiales, lo cual es erróneo. Tratan con los problemas más antiguos y fundamentales del tema, $Ax = b$ y $Ax = \lambda x$, aunque cambian y mejoran continuamente. En análisis numérico hay un remanente de la prueba de ajuste, y se desea describir algunos conceptos que han prevalecido. Éstos pueden clasificarse en tres grupos:

1. Técnicas para resolver $Ax = b$. La eliminación es un algoritmo perfecto, excepto cuando el problema en cuestión tiene propiedades especiales, como ocurre con casi todos los problemas. En la sección 7.4 la atención se centrará en la propiedad de *ser ralo(a)*, cuando la mayor parte de los elementos en A son cero. Para resolver $Ax = b$ se ***desarrollan métodos iterativos en vez de directos***. Un método iterativo se "corrige a sí mismo", y nunca llega a la respuesta exacta. El objeto es aproximarse más rápidamente que la eliminación. En algunos problemas esto es posible; en muchos otros, la eliminación es más segura y rápida si aprovecha los ceros. La competencia está lejos de terminar, y se identificará el *radio espectral* que controla la velocidad de convergencia a $x = A^{-1}b$.

2. Técnicas para resolver $Ax = \lambda x$. El problema de valores característicos es uno de los extraordinarios éxitos del análisis numérico. Está definido claramente, su importancia es evidente, aunque hasta no hace mucho nadie sabía cómo resolverlo. Se han sugerido docenas de algoritmos, y todo depende del tamaño y de las propiedades de A (y del número de valores característicos que se quieren). El lector puede solicitar LAPACK, una subrutina para valores característicos, sin conocer su contenido, aunque es mejor conocerlo. Hemos escogido dos o tres conceptos que han sustituido a casi todos sus predecesores: ***el algoritmo QR***, la familia de "***métodos de potencias***" y el reprocesamiento de una matriz simétrica para hacerla ***tridiagonal***.

Los dos primeros métodos son iterativos, y el último es directo. Hace su trabajo en un número finito de pasos, aunque no termina con los valores característicos en sí. Esto produce una matriz mucho más simple de utilizar en los pasos iterativos.

3. El número de condición de una matriz. En la sección 7.2 se intenta medir la "sensibilidad" de un problema: si A y B se modifican ligeramente, ¿cuán grande es el efecto sobre $x = A^{-1}b$? Antes de abordar esta pregunta, se requiere un método para medir A y el cambio ΔA. La longitud de un vector ya se ha definido, y ahora se necesita la ***norma de una matriz***. Luego el ***número de condición,*** y la sensibilidad de A se concluyen al multiplicar las normas de A y A^{-1}. *Las matrices de este capítulo son cuadradas.*

7.2 NORMA DE UNA MATRIZ Y NÚMERO DE CONDICIÓN

Un error y una metedura de pata son cosas muy distintas. Un error es una pequeña equivocación, quizá inevitable incluso para un matemático perfecto o una computadora perfecta. Una metedura de pata es mucho más seria, y por lo menos un orden de magnitud mayor. Cuando la computadora redondea un número después de 16 bits, eso es un error. Pero cuando un problema es tan espantosamente sensible que este error de redondeo cambia por completo la solución, entonces casi con toda seguridad alguien ha metido la pata. Nuestro objetivo en esta sección es analizar el efecto de los errores con la finalidad de evitar las meteduras de pata.

En realidad, se está continuando un análisis que empezó en el capítulo 1 con

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1.0001 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0.0001 & 1 \\ 1 & 1 \end{bmatrix}.$$

Se afirmó que B está bien acondicionada, y no es particularmente sensible al redondeo, excepto que si la eliminación gaussiana se aplica de manera imprudente, entonces la matriz se vuelve completamente vulnerable. Una metedura de pata es aceptar a 0.0001 como el primer pivote, por lo que es necesario insistir en una elección más grande y segura, mediante un intercambio de renglones de B. Cuando "pivoteo parcial" se integra al algoritmo de eliminación, la computadora busca automáticamente los *pivotes más grandes*. Así, la resistencia natural al error por redondeo deja de estar comprometida.

¿Cómo se mide esta resistencia natural, y se decide si una matriz está bien acondicionada o mal acondicionada? Si hay un cambio ligero en b o en A, ¿cuán grande es el cambio que esto produce en la solución x?

Se empieza con *un cambio en el miembro derecho*, de b a $b + \delta b$. Este error podría deberse a los datos experimentales o al redondeo. Puede suponerse que δb es pequeño, pero su dirección está fuera de nuestro control. La solución ha cambiado de x a $x + \delta x$:

Ecuación de error $A(x + \delta x) = b + \delta b$, de modo que, por sustracción, $A(\delta x) = \delta b$. (1)

Un error δb conduce a $\delta x = A^{-1}\delta b$. En la solución x hay un gran cambio cuando A^{-1} es grande, A es casi singular. El cambio en x es especialmente grande cuando δb apunta en la dirección que es más amplificada por A^{-1}.

Suponga que A es simétrica y que sus valores característicos son positivos: $0 < \lambda_1 \leq \cdots \leq \lambda_n$. Cualquier vector δb es una combinación de los vectores característicos unitarios correspondientes $x_1, \ldots, x_n$. El peor error δx, que proviene de A^{-1}, es en la dirección del primer vector característico x_1:

$$\textbf{Peor error} \qquad \text{Si} \quad \delta b = \epsilon x_1, \quad \text{entonces} \quad \delta x = \frac{\delta b}{\lambda_1}. \qquad (2)$$

El error $\|\delta b\|$ ***es amplificado por*** $1/\lambda_1$, ***que es el mayor valor característico de*** A^{-1}. Esta amplificación es máxima cuando λ_1 está próximo a cero, y A es *casi singular*.

Medir la sensibilidad completamente, por medio de λ_1 tiene una seria desventaja. Suponga que todos los elementos de A se multiplican por 1000. Así, λ_1 se multiplica por 1000 y la matriz se verá mucho menos singular. Esto ofende nuestro sentido del juego justo; un reescalamiento tan simple no puede corregir a una matriz mal acondicionada. Es cierto que δx será 1000 veces menor, pero lo mismo ocurre a la solución $x = A^{-1}b$. El error relativo $\|\delta x\|/\|x\|$ será el mismo. Al dividir entre $\|x\|$, el problema vuelve a normalizarse contra un cambio trivial de escala. Al mismo tiempo hay una normalización para δb. Nuestro problema es comparar el ***cambio relativo*** $\|\delta b\|/\|b\|$ con el ***error relativo*** $\|\delta x\|/\|x\|$.

El peor de los casos es cuando $\|\delta x\|$ es grande —con δb en la dirección del vector característico x_1—, y cuando $\|x\|$ es pequeña. La verdadera solución x debe ser lo más pequeña posible en comparación con la b verdadera. Esto significa que *el problema original* $Ax = b$ *debe estar en el otro extremo*, en la dirección del último vector característico x_n: si $b = x_n$, entonces $x = A^{-1}b = b/\lambda_n$.

Es esta combinación, $b = x_n$, y $\delta b = \epsilon x_1$, lo que hace al error relativo lo más grande posible. Estos son los casos extremos en las siguientes desigualdades:

7A Para una matriz positiva definida, la solución $x = A^{-1}b$ y el error $\delta x = A^{-1}\delta b$ siempre satisfacen

$$\|x\| \geq \frac{\|b\|}{\lambda_{\text{máx}}} \quad \text{y} \quad \|\delta x\| \leq \frac{\|\delta b\|}{\lambda_{\text{mín}}} \quad \text{y} \quad \frac{\|\delta x\|}{\|x\|} \leq \frac{\lambda_{\text{máx}}}{\lambda_{\text{mín}}}\frac{\|\delta b\|}{\|b\|}. \qquad (3)$$

La razón $c = \lambda_{\text{máx}}/\lambda_{\text{mín}}$ es el ***número de condición*** de una matriz positiva definida A.

Ejemplo 1 Los valores característicos de A son aproximadamente $\lambda_1 = 10^{-4}/2$ y $\lambda_2 = 2$:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1.0001 \end{bmatrix} \quad \text{tiene número de condición aproximado de} \quad c = 4 \cdot 10^4.$$

Debe esperarse un cambio brusco en la solución debido a los cambios normales en los datos. En el capítulo 1 se compararon las ecuaciones $Ax = b$ y $Ax' = b'$:

$$\begin{aligned} u + \quad v &= 2 \\ u + 1.0001v &= 2 \end{aligned} \qquad \begin{aligned} u + \quad v &= 2 \\ u + 1.0001v &= 2.0001. \end{aligned}$$

Los miembros derechos sólo han cambiado por $\|\delta b\| = 0.0001 = 10^{-4}$. Al mismo tiempo, la solución va de $u = 2$, $v = 0$ a $u = v = 1$. Este es el error relativo de

$$\frac{\|\delta x\|}{\|x\|} = \frac{\|(-1, 1)\|}{\|(2, 0)\|} = \frac{\sqrt{2}}{2}, \quad \text{que es igual a} \quad 2 \cdot 10^4 \frac{\|\delta b\|}{\|b\|}.$$

Sin haber hecho ninguna elección especial de la perturbación, en la solución se observa un cambio relativamente grande. Las x y δb forman ángulos de 45° con los peores casos, lo cual explica el 2 faltante entre $2 \cdot 10^4$ y la posibilidad extrema $c = 4 \cdot 10^4$.

Si $A = I$ o incluso si $A = I/10$, su número de condición es $c = \lambda_{\text{máx}}/\lambda_{\text{mín}} = 1$. Por comparación, *el determinante es una medida terrible del mal acondicionamiento*. Depende no sólo del escalamiento, sino también del orden n: si $A = I/10$, entonces el determinante de A es 10^{-n}. De hecho, esta matriz "casi singular" está lo mejor acondicionada posible.

Ejemplo 2 La matriz diferencia finita A de n por n tiene $\lambda_{\text{máx}} \approx 4$ y $\lambda_{\text{mín}} \approx \pi^2/n^2$:

$$A = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & \cdot & \\ & & \cdot & \cdot & -1 \\ & & & -1 & 2 \end{bmatrix}.$$

El número de condición es aproximadamente $c(A) = \frac{1}{2}n^2$, y esta vez la dependencia respecto al orden n es genuina. Mientras mejor se aproxime a $-u'' = f$, incrementando el número de incógnitas, más difícil es calcular la aproximación. En cierto punto de intersección, un incremento en n produce en realidad una respuesta más deficiente.

Afortunadamente para el ingeniero, esta intersección ocurre donde la precisión es bastante buena. Al trabajar con precisión simple, una computadora típica podría hacer errores por redondeo del orden de 10^{-9}. Con $n = 100$ incógnitas y $c = 5000$, el error se amplifica cuando mucho, de modo que es del orden de 10^{-5}, lo cual sigue siendo más preciso que cualquier medición normal. Sin embargo, con 10 000 incógnitas se tiene un problema, o con una aproximación 1, −4, 6, −4, 1 a $d^4u/dx^4 = f(x)$, para la cual el número de condición crece con n^4.*

Matrices no simétricas

Hasta el momento, nuestro análisis ha sido aplicado a matrices simétricas con valores característicos positivos. La hipótesis de positividad puede abandonarse fácilmente, y usar sólo valores absolutos $|\lambda|$. Pero para ir más allá de la simetría, como ciertamente se quiere, debe haber un cambio fundamental. Esto puede verse fácilmente para las matrices bastante no simétricas

$$A = \begin{bmatrix} 1 & 100 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad A^{-1} = \begin{bmatrix} 1 & -100 \\ 0 & 1 \end{bmatrix}. \tag{4}$$

Todos los valores característicos son iguales a la unidad, aunque el número de condición idóneo *no es* $\lambda_{\text{máx}}/\lambda_{\text{mín}} = 1$. El cambio relativo en x *no está* acotado por el cambio relativo en b. Compare

$$x = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ cuando } b = \begin{bmatrix} 100 \\ 1 \end{bmatrix}; \quad x' = \begin{bmatrix} 100 \\ 0 \end{bmatrix} \text{ cuando } b' = \begin{bmatrix} 100 \\ 0 \end{bmatrix}.$$

Un cambio de 1% en b ha producido un cambio centenario en x; el factor de amplificación es 100^2. Debido a que c representa una cota superior, el número de condición debe ser por lo menos 10 000. La dificultad aquí es, que un gran elemento fuera de la diagonal en A significa un elemento igualmente grande en A^{-1}. A menudo es erróneo esperar que A^{-1} se haga más pequeña a medida que A se hace más grande.

Para una definición idónea del número de condición, se considera nuevamente la ecuación (3). Se está intentando hacer pequeño a x y grande a $b = Ax$. Cuando A no es simétrica, *el máximo de* $\|Ax\|/\|x\|$ *puede encontrarse en un vector x que no es uno de los vectores característicos*. Este máximo es una excelente medida del tamaño de A. Es la ***norma*** de A.

* La regla empírica, comprobada experimentalmente, es que la computadora puede perder log c cifras decimales debido a los errores de redondeo en la eliminación gaussiana.

7B La ***norma*** de A es el número $\|A\|$ definido por

$$\|A\| = \text{máx}_{x\neq 0} \frac{\|Ax\|}{\|x\|}. \tag{5}$$

En otras palabras, $\|A\|$ acota al "poder de amplificación" de la matriz:

$$\|Ax\| \leq \|A\|\|x\| \quad \text{para todos los vectores } x. \tag{6}$$

Las matrices A y A^{-1} en la ecuación (4) tienen normas ubicadas en alguna parte entre 100 y 101. Es posible calcularlas exactamente, pero primero se desea completar la relación entre las normas y los números de condición. Debido a que $b = Ax$ y $\delta x = A^{-1}\delta b$, la ecuación (6) proporciona

$$\|b\| \leq \|A\|\|x\| \quad \text{y} \quad \|\delta x\| \leq \|A^{-1}\|\|\delta b\|. \tag{7}$$

Esta es la sustitución de la ecuación (3), cuando A no es simétrica. En el caso en que es simétrica, $\|A\|$ es la misma que $\lambda_{\text{máx}}$ y $\|A^{-1}\|$ es lo mismo que $1/\lambda_{\text{mín}}$. *La sustitución correcta para* $\lambda_{\text{máx}}/\lambda_{\text{mín}}$ *es el producto* $\|A\|\|A^{-1}\|$, que es el número de condición.

7C El ***número de condición*** de A es $c = \|A\|\|A^{-1}\|$. El error relativo satisface

δx a partir de δb $\quad \dfrac{\|\delta x\|}{\|x\|} \leq c\dfrac{\|\delta b\|}{\|b\|}$ directamente de la ecuación (7). (8)

Si se perturba la matriz A en vez del miembro derecho b, entonces

δx a partir de δA $\quad \dfrac{\|\delta x\|}{\|x+\delta x\|} \leq c\dfrac{\|\delta A\|}{\|A\|}$ directamente de la ecuación (10) que se muestra a continuación. (9)

Lo que es extraordinario es que el mismo número de condición aparece en la ecuación (9), cuando se perturba la matriz misma: Si $Ax = b$ y $(A + \delta A)(x + \delta x) = b$, entonces por sustracción

$$A\delta x + \delta A(x + \delta x) = 0, \quad \text{o} \quad \delta x = -A^{-1}(\delta A)(x + \delta x).$$

Al multiplicar por δA se amplifica un vector por no más de $\|\delta A\|$, y al multiplicar por A^{-1} se amplifica por no más de $\|A^{-1}\|$. Luego, $\|\delta x\| < \|A^{-1}\|\|\delta A\|\|x + \delta x\|$, que es

$$\frac{\|\delta x\|}{\|x + \delta x\|} \leq \|A^{-1}\|\|\delta A\| = c\frac{\|\delta A\|}{\|A\|}. \tag{10}$$

Estas desigualdades significan que el error por redondeo proviene de dos fuentes. Una es la *sensibilidad natural* del problema, medida por c. La otra es el error verdadero δb o δA. Esto constituyó la base del análisis de error de Wilkinson. Debido a que la eliminación realmente produce factores aproximados L' y U', resuelve la ecuación con la matriz errónea $A + \delta A = L'U'$, en vez de hacerlo con la matriz correcta $A = LU$. Wilkinson demostró que el pivoteo parcial controla δA, de modo que *la carga del error por redondeo es llevada por el número de condición c*.

Una fórmula para la norma

La norma de A mide la mayor cantidad por la cual cualquier vector (vector característico o no) es amplificado por la matriz multiplicación: $\|A\| = \text{máx}\,(\|Ax\|/\|x\|)$. La norma de la

matriz identidad es 1. Para calcular la norma, ambos miembros se elevan al cuadrado para llegar a la A^TA simétrica:

$$\|A\|^2 = \text{máx}\, \frac{\|Ax\|^2}{\|x\|^2} = \text{máx}\, \frac{x^TA^TAx}{x^Tx}. \tag{11}$$

7D $\|A\|$ es la raíz cuadrada del mayor valor característico de A^TA: $\|A\|^2 = \lambda_{\text{máx}}(A^TA)$. El vector que A amplifica más, es el vector característico correspondiente de A^TA:

$$\frac{x^TA^TAx}{x^Tx} = \frac{x^T(\lambda_{\text{máx}}x)}{x^Tx} = \lambda_{\text{máx}}(A^TA) = \|A\|^2. \tag{12}$$

En la figura 7.1 se muestra una matriz no simétrica con valores característicos $\lambda_1 = \lambda_2 = 1$ y norma $\|A\| = 1.618$. En este caso A^{-1} tiene la misma norma. Los puntos más alejado y más próximo Ax sobre la elipse provienen de los vectores característicos de A^TA, no de A.

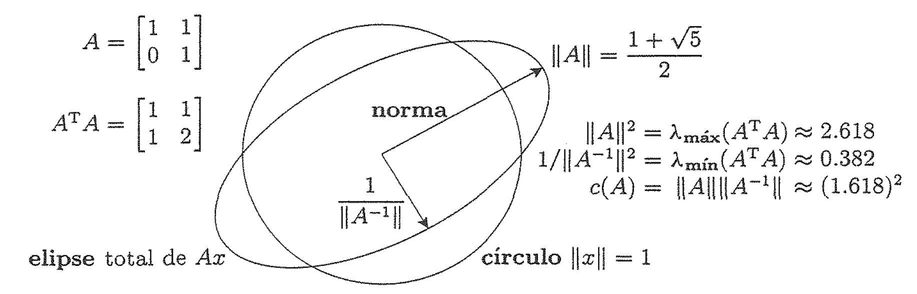


Figura 7.1 Las normas de A y A^{-1} provienen de la Ax más larga y más corta.

Nota 1 La norma y el número de condición no se calculan realmente en la práctica, sino que sólo se estiman. No hay tiempo para resolver un problema de valores característicos para $\lambda_{\text{máx}}(A^TA)$.

Nota 2 En la ecuación por mínimos cuadrados $A^TAx = A^Tb$, el número de condición $c(A^TA)$ es el *cuadrado* de $c(A)$. Al formar A^TA, un problema sano puede convertirse en uno enfermo. Quizá sea necesario ortogonalizar A con el proceso de Gram-Schmidt, en vez de calcularla con A^TA.

Nota 3 Los ***valores singulares*** de A en la DVS son *las raíces cuadradas de los valores característicos de* A^TA. Por la ecuación (12), otra fórmula para la norma es $\|A\| = \sigma_{\text{máx}}$. Las matrices ortogonales U y V dejan sin cambio las longitudes en $\|Ax\| = \|U\Sigma V^Tx\|$. Así, el mayor $\|Ax\|/\|x\|$ proviene del mayor σ en la matriz diagonal Σ.

Nota 4 El error por redondeo también entra en $Ax = \lambda x$. ¿Cuál es el número de condición del problema de valores característicos? *El número de condición de la matriz de diagonalización S mide la sensibilidad de los valores característicos.* Si μ es un valor característico de $A + E$, entonces su distancia a uno de los valores característicos de A es

$$|\mu - \lambda| \le \|S\|\,\|S^{-1}\|\,\|E\| = c(S)\|E\|. \tag{13}$$

Con vectores característicos ortonormales y $S = Q$, el problema de valores característicos está perfectamente acondicionado: $c(Q) = 1$. El cambio $\delta\lambda$ en los valores característicos no es mayor que el cambio δA. En consecuencia, el mejor de los casos es cuando A es simétrica, o más generalmente, cuando $AA^T = A^TA$. Entonces A es una matriz normal; su S diagonalizadora es una Q ortogonal (véase la sección 5.6).

Si x_k es la k-ésima columna de S y y_k es el k-ésimo renglón de S^{-1}, entonces λ_k cambia por

$$\delta\lambda_k = y_kEx_k + \text{términos de orden } \|E\|^2. \tag{14}$$

En la práctica, y_kEx_k es una estimación realista de $\delta\lambda$. La idea en todo buen algoritmo es mantener la matriz error E lo más pequeña posible —usualmente al insistir, como se hará en la siguiente sección, en matrices ortogonales en cada paso del cálculo de λ.

Conjunto de problemas 7.2

1. Para una matriz ortogonal Q, demuestre que $\|Q\| = 1$ y también que $c(Q) = 1$. Las matrices ortogonales (y sus múltiplos αQ) son las únicas matrices perfectamente acondicionadas.
2. ¿Qué "famosa" desigualdad proporciona $\|(A + B)x\| \le \|Ax\| + \|Bx\|$, y por qué se concluye de la ecuación (5) que $\|A + B\| \le \|A\| + \|B\|$?
3. Explique por qué $\|ABx\| \le \|A\|\,\|B\|\,\|x\|$, y concluya de la ecuación (5) que $\|AB\| \le \|A\|\|B\|$. Demuestre que esto también implica $c(AB) \le c(A)c(B)$.
4. Para la matriz positiva definida $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$, calcule $\|A^{-1}\| = 1/\lambda_1$, $\|A\| = \lambda_2$, y $c(A) = \lambda_2/\lambda_1$. Para un miembro derecho b y una perturbación δb de modo que el error sea el peor posible, $\|\delta x\|/\|x\| = c\|\delta b\|/\|b\|$.
5. Demuestre que si λ es cualquier valor característico de A, $Ax = \lambda x$, entonces $\|\lambda\| \le \|A\|$.
6. Las matrices en la ecuación (4) tienen normas entre 100 y 101. ¿Por qué?
7. Compare los valores característicos de A^TA y AA^T, para demostrar que $\|A\| = \|A^T\|$.
8. Para una matriz positiva definida A, la descomposición de Cholesky es $A = LDL^T = R^TR$, donde $R = \sqrt{D}L^T$. A partir de la ecuación (12), demuestre directamente que el número de condición de $c(R)$ es la raíz cuadrada de $c(A)$. La eliminación sin intercambios de renglones no puede afectar a una matriz positiva definida, ya que $c(A) = c(R^T)c(R)$.
9. Demuestre que máx$|\lambda|$ no es una norma verdadera, encontrando contraejemplos de 2 por 2 para $\lambda_{\text{máx}}(A + B) \le \lambda_{\text{máx}}(A) + \lambda_{\text{máx}}(B)$ y $\lambda_{\text{máx}}(AB) \le \lambda_{\text{máx}}(A)\lambda_{\text{máx}}(B)$.
10. Demuestre que los valores característicos de $B = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$ son $\pm\sigma_i$, los valores singulares de A. *Sugerencia*: Intente B^2.
11. *a*) A y A^{-1}, ¿tienen el mismo número de condición c?
    b) En paralelo con la cota superior (8) sobre el error, demuestre una cota inferior:

$$\frac{\|\delta x\|}{\|x\|} \ge \frac{1}{c}\frac{\|\delta b\|}{\|b\|}. \quad \text{(Considere } A^{-1}b = x \text{ en vez de } Ax = b.)$$

12. Encuentre las normas $\lambda_{\text{máx}}$ y los números de condición $\lambda_{\text{máx}}/\lambda_{\text{mín}}$ de las siguientes matrices positivas definidas:

$$\begin{bmatrix} 100 & 0 \\ 0 & 2 \end{bmatrix} \quad \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}.$$

13. Encuentre las normas y los números de condición, a partir de las raíces cuadradas de $\lambda_{\text{máx}}(A^{\mathrm{T}}A)$ y $\lambda_{\text{mín}}(A^{\mathrm{T}}A)$:

$$\begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

14. Demuestre que el número de condición $\|A\|\|A^{-1}\|$ es por lo menos 1.

15. ¿Por qué I es la única matriz positiva definida simétrica que cumple $\lambda_{\text{máx}} = \lambda_{\text{mín}} = 1$? Así, las únicas matrices con $\|A\| = 1$ y $\|A^{-1}\| = 1$ deben cumplir $A^{\mathrm{T}}A = I$. Son matrices ______.

16. Las matrices ortogonales tienen norma $\|Q\| = 1$. Si $A = QR$, demuestre que $\|A\| \leq \|R\|$ y también que $\|R\| \leq \|A\|$. Así, $\|A\| = \|Q\|\|R\|$. Encuentre un ejemplo de $A = LU$ con $\|A\| < \|L\|\|U\|$.

17. (Sugerido por Moler y Van Loan) Calcule $b - Ay$ y $b - Az$ cuando

$$b = \begin{bmatrix} 0.217 \\ 0.254 \end{bmatrix} \quad A = \begin{bmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{bmatrix} \quad y = \begin{bmatrix} 0.341 \\ -0.087 \end{bmatrix} \quad z = \begin{bmatrix} 0.999 \\ -1.0 \end{bmatrix}.$$

Para resolver $Ax = b$, ¿y está más próximo que z? Conteste en dos formas: Compare el *residuo* $b - Ay$ con $b - Az$. Luego compare y y z con el verdadero $x = (1, -1)$. Algunas veces se desea un residuo pequeño, y algunas veces, un δx pequeño.

Los problemas 18 a 20 son sobre normas vectoriales distintas a la norma de costumbre $\|x\| = \sqrt{x \cdot x}$.

18. La "norma ℓ^1" es $\|x\|_1 = |x_1| + \cdots + |x_n|$. La "norma ℓ^∞" es $\|x\|_\infty = \text{máx } |x_i|$. Calcule $\|x\|$, $\|x\|_1$, y $\|x\|_\infty$ para los vectores

$$x = (1, 1, 1, 1, 1) \quad \text{y} \quad x = (0.1, 0.7, 0.3, 0.4, 0.5).$$

19. Demuestre que $\|x\|_\infty \leq \|x\| \leq \|x\|_1$. A partir de la desigualdad de Schwarz, demuestre que las razones $\|x\|/\|x\|_\infty$ y $\|x\|_1/\|x\|$ nunca son mayores que $\sqrt{n}$. ¿Con qué vector $(x_1, \ldots, x_n)$ se obtienen razones iguales a $\sqrt{n}$?

20. Todas las normas vectoriales deben cumplir la *desigualdad del triángulo*. Demuestre que

$$\|x + y\|_\infty \leq \|x\|_\infty + \|y\|_\infty \quad \text{y} \quad \|x + y\|_1 \leq \|x\|_1 + \|y\|_1.$$

21. Por eliminación, calcule la inversa exacta de la matriz de Hilbert A. Luego, calcule de nuevo A^{-1} redondeando todos los números hasta tres cifras:

$$\text{En MATLAB: } A = \text{hilb}(3) = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}.$$

22. Para la misma A, calcule $b = Ax$ para $x = (1, 1, 1)$ y $x = (0, 6, -3.6)$. Un cambio pequeño Δb produce un gran cambio Δx.

23. Calcule $\lambda_{\text{máx}}$ y $\lambda_{\text{mín}}$ para la matriz de Hilbert de 8 por 8 $a_{ij} = 1/(i + j - 1)$. Si $Ax = b$ con $\|b\| = 1$, ¿cuán grande puede ser $\|x\|$? Si el error por redondeo de b es menor que 10^{-16}, ¿cuán grande es el error que esto puede provocar en x?

24. Si se conocen L, U, Q, y R, ¿qué es más rápido resolver, $LUx = b$ o $QRx = b$?

25. Escoja el pivote más grande disponible en cada columna (pivoteo parcial), para factorizar cada A en $PA = LU$:

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 0 \\ 0 & 2 & 0 \end{bmatrix}.$$

26. Encuentre la factorización LU de $A = \begin{bmatrix} \epsilon & 1 \\ 1 & 1 \end{bmatrix}$. En su computadora, resuelva por eliminación cuando $\epsilon = 10^{-3}, 10^{-6}, 10^{-9}, 10^{-12}, 10^{-15}$:

$$\begin{bmatrix} \epsilon & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 + \epsilon \\ 2 \end{bmatrix}.$$

El x verdadero es $(1, 1)$. Elabore una tabla para mostrar el error para cada ϵ. Intercambie las dos ecuaciones y vuelva a resolver: los errores casi deben desaparecer.

7.3 CÁLCULO DE VALORES CARACTERÍSTICOS

No hay ninguna forma mejor que otra para encontrar los valores característicos de una matriz. Sin embargo, ciertamente existen algunas formas terribles que nunca deben intentarse, así como también algunas ideas que merecen un sitio permanente. Empezamos con la descripción de un método bastante fácil y aproximado: el ***método de potencias***, cuyas propiedades de convergencia son fáciles de comprender. Para demostrar este método en acción, en la página del curso *web.mit.edu/18.06* se agregó una animación gráfica (sonora).

De manera continua nos dirigimos hacia otro algoritmo más complicado, que empieza por hacer tridiagonal a una matriz simétrica y termina por hacerla virtualmente diagonal. Este segundo paso se realiza repitiendo el proceso de Gram-Schmidt, por lo que se denomina ***método QR***.

El método de potencias normal opera, con base en el principio de una ecuación en diferencias. Empieza con una conjetura inicial u_0 y sucesivamente forma $u_1 = Au_0$, $u_2 = Au_1$ y en general $u_{k+1} = Au_k$. Cada paso es una multiplicación matriz-vector. Luego de k pasos produce $u_k = A^k u_0$, aunque la matriz A^k nunca aparece. La cuestión esencial es que la multiplicación por A debe ser fácil —si la matriz es grande, mejor sería si fuese rala— debido a que la convergencia al vector característico a menudo es muy lenta. Suponiendo que A tiene un conjunto completo de vectores característicos $x_1, \ldots, x_n$, el vector u_k está dado por la fórmula de costumbre:

Vectores característicos ponderados por λ^k $\quad u_k = c_1\lambda_1^k x_1 + \cdots + c_n\lambda_n^k x_n.$

Suponga que el valor característico más grande λ_n es todo en sí: no hay otro valor característico de la misma magnitud, y $|\lambda_1| \leq \cdots \leq |\lambda_{n-1}| < |\lambda_n|$. Entonces, a medida que la conjetura inicial u_0 contenía *alguna* componente x_n del vector característico, de modo que $c_n \neq 0$, esta componente domina gradualmente en u_k:

$$\frac{u_k}{\lambda_n^k} = c_1\left(\frac{\lambda_1}{\lambda_n}\right)^k x_1 + \cdots + c_{n-1}\left(\frac{\lambda_{n-1}}{\lambda_n}\right)^k x_{n-1} + c_n x_n. \tag{1}$$

Los vectores u_k apuntan cada vez con mayor precisión hacia la dirección de x_n. Su factor de convergencia es la razón $r = |\lambda_{n-1}|/|\lambda_n|$. Es justo como la convergencia hacia un estado estacionario, para una matriz de Markov, excepto que ahora λ_n puede no ser igual a 1. El factor de escalamiento λ_n^k en la ecuación (1) evita que u_k se vuelva muy grande o muy pequeño, en caso de que $|\lambda_n| > 1$ o $|\lambda_n| < 1$.

A menudo apenas es posible dividir cada u_k entre su primera componente α_k antes de continuar con el paso siguiente. Con este simple escalamiento, el método de potencias $u_{k+1} = Au_k/\alpha_k$ converge a un múltiplo de x_n. ***Los factores de escalamiento α_k tienden a λ_n***

Ejemplo 1 El u_k tiende al vector característico $\begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 0.667 \\ 0.333 \end{bmatrix}$ cuando $A = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}$ es la matriz de movimientos de población en la sección 1.3:

$$u_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad u_1 = \begin{bmatrix} 0.9 \\ 0.1 \end{bmatrix}, \quad u_2 = \begin{bmatrix} 0.83 \\ 0.17 \end{bmatrix}, \quad u_3 = \begin{bmatrix} 0.781 \\ 0.219 \end{bmatrix}, \quad u_4 = \begin{bmatrix} 0.747 \\ 0.253 \end{bmatrix}.$$

Si $r = |\lambda_{n-1}|/|\lambda_n|$ es próximo a 1, entonces la convergencia es muy lenta. En muchas aplicaciones $r > 0.9$, lo cual significa que para alcanzar un dígito más se requieren más de 20 iteraciones. (En el ejemplo se tenía $r = 0.7$, y aun así era lenta). Si $r = 1$, lo que significa $|\lambda_{n-1}| = |\lambda_n|$, entonces es probable que la convergencia ni siquiera ocurra. Esto sucede (en el paquete sonoro) para un par conjugado complejo $\lambda_{n-1} = \bar{\lambda}_n$. Hay varias formas de evitar esta limitación. A continuación se describen tres.

1. El ***método de potencias en bloque*** funciona con varios vectores a la vez, en lugar de u_k. Si p vectores ortonormales se multiplican por A, y luego para ortogonalizarlos se aplica de nuevo el proceso de Gram-Schmidt —lo cual es un solo paso del método— entonces la razón de convergencia se convierte en $r' = |\lambda_{n-p}|/|\lambda_n|$. Se obtendrán aproximaciones para p valores característicos distintos y sus vectores característicos.
2. El ***método de potencias inverso*** opera con A^{-1} en vez de A. Un simple paso es $v_{k+1} = A^{-1}v_k$, lo cual significa que se resuelve el sistema $Av_{k+1} = v_k$ (¡y se ahorran los factores L y U!) En este caso se converge al *valor característico más pequeño* λ_1 y su vector característico x_1, en el supuesto de que $|\lambda_1| < |\lambda_2|$. A menudo, lo que se requiere en las aplicaciones es λ_1, y entonces la iteración inversa es una opción automática.
3. El ***método de potencias inverso desplazado*** es el mejor de todos. A se sustituye por $A - \alpha I$. Cada valor característico es desplazado por α, y el factor de convergencia para el método inverso cambia a $r'' = |\lambda_1 - \alpha|/|\lambda_2 - \alpha|$. Si α es una aproximación aceptable a λ_1, entonces r'' es muy pequeño y la convergencia se acelera enormemente. Cada paso del método resuelve $(A - \alpha I)w_{k+1} = w_k$:
$$w_k = \frac{c_1x_1}{(\lambda_1 - \alpha)^k} + \frac{c_2x_2}{(\lambda_2 - \alpha)^k} + \cdots + \frac{c_nx_n}{(\lambda_n - \alpha)^k}.$$
Cuando α está próximo a λ_1, el primer término domina sólo al cabo de uno o dos pasos. Si λ_1 ya se ha calculado con algún otro algoritmo (como QR), entonces α es este valor calculado. Un procedimiento normal es factorizar $A - \alpha I$ en LU y resolver $Ux_1 = (1, 1, \ldots, 1)$ por sustitución hacia atrás.

Si λ_1 no se ha aproximado aún, entonces el método de potencias inverso desplazado debe generar su propia elección de α. Es posible hacer variar $\alpha = \alpha_k$ en cada paso si así se desea, de modo que $(A - \alpha_k I)w_{k+1} = w_k$.

Cuando A es simétrica, una opción bastante precisa es el ***cociente de Rayleigh***:

$$\text{desplazado por} \qquad \alpha_k = R(w_k) = \frac{w_k^{\mathrm{T}}Aw_k}{w_k^{\mathrm{T}}w_k}.$$

Este cociente $R(x)$ tiene un mínimo en el vector característico verdadero x_1. Su gráfica es como el fondo de una parábola, de modo que el error $\lambda_1 = \alpha_k$ es aproximadamente el cuadrado del error del vector característico. Los factores de convergencia $|\lambda_1 - \alpha_k|/|\lambda_2 - \alpha_k|$ en sí convergen a cero. Así, estos desplazamientos de cocientes de Rayleigh proporcionan una *convergencia cúbica* de α_k a λ_1.*

Formas tridiagonales y de Hessenberg

El método de potencias es razonable sólo para una matriz grande y rala. Cuando demasiados elementos son distintos de cero, este método es un error. En consecuencia, se plantea la pregunta de si existe alguna forma sencilla de *crear ceros*. Este es el objetivo de los párrafos siguientes.

Puede afirmarse que después de calcular una matriz semejante $Q^{-1}AQ$ con más ceros que A, no se intenta regresar al método de potencias. Hay variantes mucho más poderosas, y parece que la mejor es el algoritmo QR. (El método de potencias inverso desplazado tiene su sitio al final, para encontrar el vector característico.) El primer paso es producir rápidamente tantos ceros como sea posible, usando una matriz ortogonal Q. Si A es simétrica, entonces también lo es $Q^{-1}AQ$. Ningún elemento puede hacerse peligrosamente grande porque Q preserva la longitud.

Para pasar de A a $Q^{-1}AQ$, hay dos posibilidades principales: Es posible producir un cero en cada paso (como en la eliminación), o puede trabajarse con toda una columna a la vez. Para un simple cero, es fácil utilizar una rotación del plano como se ilustra en la ecuación (7), que se encuentra cerca del final de esta sección, que tenga a cos θ y sen θ en un bloque de 2 por 2. Luego, puede ciclarse a través de todos los elementos por debajo de la diagonal, escogiendo en cada paso una rotación θ que produzca un cero; este es el ***método de Jacobi***. Fracasa en diagonalizar A después de un número finito de rotaciones, ya que los ceros de pasos previos se destruyen cuando se crean ceros posteriores.

Para preservar los ceros y detenerse, es necesario plantear lo anterior para menos que una forma triangular. ***La forma de Hessenberg acepta una diagonal distinta de cero por abajo de la diagonal principal.*** Si una matriz de Hessenberg es simétrica, entonces sólo tiene tres diagonales distintas de cero.

Una serie de rotaciones en los planos idóneos produce los ceros requeridos. Householder encontró una nueva manera de obtener exactamente lo mismo. Una ***transformación de Householder*** es una matriz reflexión, determinada por un vector v:

$$\textbf{Matriz de Householder} \qquad H = I - 2\frac{vv^{\mathrm{T}}}{\|v\|^2}.$$

A menudo, v se normaliza para convertirlo en un vector unitario $u = v/\|v\|$, y entonces H se transforma en $I - 2uu^{\mathrm{T}}$.

---

*Convergencia lineal significa que cada paso multiplica el error por un factor fijo $r < 1$. Convergencia cuadrática significa que en cada paso el error se eleva al cuadrado, como en el método de Newton $x_{k+1} - x_k = -f(x_k)/f'(x_k)$ para resolver $f(x) = 0$. La convergencia cúbica requiere de 10^{-1} a 10^{-3} a 10^{-9}.

En cualquier caso, H es tanto *simétrica* como *ortogonal*:

$$H^TH = (I - 2uu^T)(I - 2uu^T) = I - 4uu^T + 4uu^Tuu^T = I.$$

Así, $H = H^T = H^{-1}$. El plan de Householder era producir ceros con estas matrices, y su éxito depende de la siguiente identidad $Hx = -\sigma z$:

7E Suponga que z es el vector columna $(1, 0, \ldots, 0)$, $\sigma = \|x\|$, y $v = x + \sigma z$. Entonces $Hx = -\sigma z = (-\sigma, 0, \ldots, 0)$. El vector Hx termina en ceros, como se quería.

La demostración consiste en calcular Hx y llegar a $-\sigma z$:

$$\begin{aligned} Hx = x - \frac{2vv^Tx}{\|v\|^2} &= x - (x + \sigma z)\frac{2(x+\sigma z)^Tx}{(x+\sigma z)^T(x+\sigma z)} \\ &= x - (x + \sigma z) \qquad (\text{porque } x^Tx = \sigma^2) \\ &= -\sigma z. \end{aligned} \tag{2}$$

Esta identidad puede utilizarse de inmediato, sobre la primera columna de A. A la $Q^{-1}AQ$ final le permite una diagonal diferente de cero abajo de la diagonal principal (forma de Hessenberg). En consecuencia, *sólo participan los elementos que están estrictamente debajo de la diagonal*:

$$x = \begin{bmatrix} a_{21} \\ a_{31} \\ \vdots \\ a_{n1} \end{bmatrix}, \quad z = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad Hx = \begin{bmatrix} -\sigma \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{3}$$

En este punto la matriz H de Householder es sólo de orden $n - 1$, de modo que está incrustada en la esquina inferior derecha de una matriz U_1 de tamaño completo:

$$U_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & & & & \\ 0 & & H & & \\ 0 & & & & \\ 0 & & & & \end{bmatrix} = U_1^{-1}, \qquad \text{y} \qquad U_1^{-1}AU_1 = \begin{bmatrix} a_{11} & * & * & * & * \\ -\sigma & * & * & * & * \\ \mathbf{0} & * & * & * & * \\ \mathbf{0} & * & * & * & * \\ \mathbf{0} & * & * & * & * \end{bmatrix}.$$

Esta primera etapa está completa, y $U_1{}^{-1}AU_1$ tiene la primera columna requerida. En la segunda etapa, x consta de los $n - 2$ últimos elementos en la segunda columna (tres estrellas en negritas). Así, H_2 es de orden $n - 2$. Cuando está incrustada en U_2, produce

$$U_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & & & \\ 0 & 0 & & H_2 & \\ 0 & 0 & & & \end{bmatrix} = U_2^{-1}, \qquad U_2^{-1}(U_1^{-1}AU_1)U_2 = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ \mathbf{0} & * & * & * & * \\ \mathbf{0} & \mathbf{0} & * & * & * \\ \mathbf{0} & \mathbf{0} & * & * & * \end{bmatrix}.$$

U_3 se encargará de la tercera columna. Para una matriz de 5 por 5, se alcanza la forma de Hessenberg (tiene seis ceros). Como regla general, Q es el producto de todas las matrices $U_1U_2 \cdots U_{n-2}$, y el número de operaciones necesarias para calcularla es de orden n^3.

Ejemplo 2 (para cambiar $a_{13} = a_{31}$ en cero)

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \qquad x = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad H = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

Al incrustar H en Q, el resultado $Q^{-1}AQ$ es tridiagonal.

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}, \qquad Q^{-1}AQ = \begin{bmatrix} 1 & -1 & \mathbf{0} \\ -1 & 0 & 1 \\ \mathbf{0} & 1 & 1 \end{bmatrix}.$$

$Q^{-1}AQ$ es una matriz que ya está lista para revelar sus valores característicos: el algoritmo QR está listo para comenzar, aunque a continuación se hace una digresión para mencionar otras dos aplicaciones de estas mismas matrices H de Householder.

1. *La factorización de Gram-Schmidt* $A = QR$. Recuerde que R ha de ser triangular superior. Ya no es necesario aceptar una diagonal distinta de cero extra abajo de la diagonal principal, ya que no hay matrices multiplicando por la derecha que estropeen los ceros. El primer paso en la construcción de Q es trabajar con toda la primera columna de A:

$$x = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \qquad z = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad v = x + \|x\|z, \qquad H_1 = I - 2\frac{vv^T}{\|v\|^2}.$$

La primera columna de H_1A es igual a $-\|x\|z$. Es cero abajo de la diagonal principal, y es la primera columna de R. El segundo paso trabaja con la segunda columna de H_1A, desde el pivote hacia abajo, y produce una H_2H_1A que es cero abajo de ese pivote. (Todo el algoritmo es como la eliminación, aunque ligeramente más lento.) El resultado de $n - 1$ pasos es una matriz triangular superior R, aunque la matriz que registra los pasos no es una triangular inferior L. En vez de lo anterior, es el producto $Q = H_1H_2 \cdots H_{n-1}$, que puede almacenarse en esta forma factorizada (sólo se preservan los vs) y nunca se calcula explícitamente. Así se completa el proceso de Gram-Schmidt.

2. *La descomposición del valor singular* $U^TAV = \Sigma$. La matriz diagonal Σ tiene la misma forma que A, y sus elementos (los valores singulares) son las raíces cuadradas de los valores característicos de A^TA. Debido a que las transformaciones de Householder sólo pueden *preparar* el problema de valores característicos, no es de esperar que produzcan Σ. En vez de ello, producen de manera estable una *matriz bidiagonal*, con ceros en todas partes, excepto a lo largo de la diagonal principal y la diagonal arriba de ésta.

El primer paso hacia la DVS es exactamente como en QR arriba: x es la primera columna de A, y H_1x es cero abajo del primer pivote. El paso siguiente es multiplicar por la derecha por una $H^{(1)}$ que produzca ceros, como se indica a lo largo del primer renglón:

$$A \to H_1A = \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix} \to H_1AH^{(1)} = \begin{bmatrix} * & * & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & * & * & * \\ \mathbf{0} & * & * & * \end{bmatrix}. \tag{4}$$

Luego, dos transformaciones de Householder alcanzan rápidamente la forma bidiagonal:

$$H_2H_1AH^{(1)} = \begin{bmatrix} * & * & 0 & 0 \\ 0 & * & * & * \\ 0 & 0 & * & * \end{bmatrix} \quad \text{y} \quad H_2H_1AH^{(1)}H^{(2)} = \begin{bmatrix} * & * & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & * & * & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & * & * \end{bmatrix}.$$

El algoritmo *QR* para calcular valores característicos

El algoritmo es casi mágicamente sencillo. Empieza con A_0, la factoriza en Q_0R_0 con el proceso de Gram-Schmidt, y luego *invierte los factores*: $A_1 = R_0Q_0$. La nueva matriz A_1 es *semejante* a la original porque $Q_0^{-1}A_0Q_0 = Q_0^{-1}(Q_0R_0)Q_0 = A_1$. Así, el proceso continúa sin cambio en los valores característicos:

$$\text{Toda } A_k \text{ es igual} \qquad A_k = Q_kR_k \quad \text{y entonces} \quad A_{k+1} = R_kQ_k. \tag{5}$$

Esta ecuación describe el *algoritmo QR no desplazado*, y casi siempre A_k tiende a una forma triangular. Sus elementos diagonales tienden a sus valores característicos, que también son los valores característicos de A_0. Si ya existía algún procesamiento para obtener la forma tridiagonal, entonces A_0 está relacionada con la A absolutamente original, mediante $Q^{-1}AQ = A_0$.

Como está, el algoritmo *QR* es bueno pero no tanto. Para hacerlo especial, requiere dos refinamientos: es necesario permitir desplazamientos a $A_k - \alpha_k I$, y debe asegurarse que la factorización *QR* en cada paso sea muy rápida.

1. El algoritmo desplazado. Si el número α_k está muy próximo de un valor característico, el paso en la ecuación (5) debe desplazarse de inmediato por α_k (lo cual cambia Q_k y R_k):

$$A_k - \alpha_k I = Q_kR_k \quad \text{y entonces} \quad A_{k+1} = R_kQ_k + \alpha_k I. \tag{6}$$

Esta matriz A_{k+1} es semejante a A_k (siempre los mismos valores característicos):

$$Q_k^{-1}A_kQ_k = Q_k^{-1}(Q_kR_k + \alpha_k I)Q_k = A_{k+1}.$$

Lo que ocurre en la práctica es que el elemento (n, n) de A_k; el que está en la esquina inferior derecha, es el primero en tender a un valor característico. Ese elemento es la elección más simple y conocida para el desplazamiento α_k. Normalmente esto produce convergencia cuadrática, y en el caso simétrico incluso convergencia cúbica, al menor valor característico. Al cabo de tres o cuatro pasos del algoritmo desplazado, la matriz A_k se ve como sigue:

$$A_k = \left[\begin{array}{ccc|c} * & * & * & * \\ * & * & * & * \\ 0 & * & * & * \\ \hline 0 & 0 & \epsilon & \lambda_1' \end{array}\right], \quad \text{con } \epsilon \ll 1.$$

El λ'_1 se acepta como una aproximación muy próxima al λ_1 verdadero. Para encontrar el siguiente valor característico, el algoritmo *QR* continúa con la matriz más pequeña (de 3 por 3, en la ilustración) en la esquina superior izquierda. Sus elementos subdiagonales se reducen algo en los primeros pasos *QR*, y otros dos pasos son suficientes para encontrar λ_2. Así se obtiene un procedimiento sistemático para encontrar todos los valores característicos. De hecho, ***el método QR ya está descrito por completo.*** Sólo queda encontrar los vec-

tores característicos, lo cual es un simple paso de potencias inverso, y usar los ceros creados por Householder.

2. Cuando A_0 es tridiagonal o de Hessenberg, cada paso *QR* es muy rápido. El proceso de Gram-Schmidt (factorización en *QR*) requiere $O(n^3)$ operaciones para una matriz completa A. Para una matriz de Hessenberg, lo anterior se vuelve en $O(n^2)$, y para una matriz tridiagonal, es $O(n)$. Por fortuna, cada nueva A_k está nuevamente en forma de Hessenberg o tridiagonal:

$$\textbf{Q_0 es de Hessenberg} \qquad Q_0 = A_0R_0^{-1} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \end{bmatrix} \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{bmatrix}.$$

Puede comprobarse fácilmente que esta multiplicación deja a Q_0 con los mismos tres ceros que A_0. *Una matriz de Hessenberg multiplicada por una matriz triangular es una matriz de Hessenberg*. Lo mismo es cierto para una matriz triangular multiplicada por una matriz de Hessenberg:

$$\textbf{A_1 es de Hessenberg} \qquad A_1 = R_0Q_0 = \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{bmatrix} \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \end{bmatrix}.$$

El caso simétrico es todavía mejor, ya que $A_1 = Q_0^{-1}A_0Q_0 = Q_0^TA_0Q_0$ permanece simétrica. Por el razonamiento que acaba de presentarse, A_1 también es de Hessenberg. Así, A_1 debe ser *tridiagonal*. Lo mismo es válido para $A_2, A_3, \ldots$, y *cada paso QR empieza con una matriz tridiagonal*.

El último punto es la factorización en sí, que produce las Q_k y R_k a partir de cada A_k (o realmente, de $A_k - \alpha_k I$). Householder puede usarse de nuevo, pero es más sencillo aniquilar cada elemento subdiagonal a la vez, mediante una "rotación del plano" P_{ij}. La primera es P_{21}:

$$\textbf{Rotación para eliminar a a_{21}} \quad P_{21}A_k = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & & \\ \operatorname{sen}\theta & \cos\theta & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} a_{11} & * & * & * \\ a_{21} & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \end{bmatrix} \tag{7}$$

El elemento (2, 1) en este producto es $a_{11} \operatorname{sen}\theta + a_{21}\cos\theta$, y se escoge el ángulo θ que hace cero esta combinación. La siguiente rotación P_{32} se elige de forma semejante, para eliminar el elemento (3, 2) de $P_{32}P_{21}A_k$. Después de $n - 1$ rotaciones, se tiene R_0:

$$\textbf{Factor triangular} \qquad R_k = P_{n\,n-1} \cdots P_{32}P_{21}A_k. \tag{8}$$

Más información acerca de este extraordinario algoritmo para cálculos científicos, puede encontrarse en libros sobre álgebra lineal numérica. Se menciona un método más: ***Arnoldi*** en ARPACK, para matrices grandes. Ortogonaliza la sucesión de Krilov $x, Ax, A^2x, \ldots$, mediante el proceso de Gram-Schmidt. Si requiere los valores característicos de una matriz grande, ¡no use $\det(A - \lambda I)$!

Conjunto de problemas 7.3

1. Para la matriz $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$, con valores característicos $\lambda_1 = 1$ y $\lambda_2 = 3$, aplique tres veces el método de potencias $u_{k+1} = Au_k$ a la conjetura inicial $u_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. ¿Cuál es el vector limitante u_∞?

2. Para la misma A y la conjetura inicial $u_0 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$, compare tres pasos de potencias inversos con un paso desplazado con $\alpha = u_0^{\mathrm{T}} A u_0 / u_0^{\mathrm{T}} u_0$:

$$u_{k+1} = A^{-1} u_k = \frac{1}{3}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} u_k \text{ o bien, } u = (A - \alpha I)^{-1} u_0.$$

El vector limitante u_∞ ahora es un múltiplo del otro vector característico (1, 1).

3. Explique por qué $|\lambda_n / \lambda_{n-1}|$ controla la convergencia del método de potencias de costumbre. Construya una matriz A para la que este método *no converge*.

4. La matriz de Markov $A = \begin{bmatrix} 0.9 & 0.3 \\ 0.1 & 0.7 \end{bmatrix}$ tiene $\lambda = 1$ y 0.6, y el método de potencias $u_k = A^k u_0$ converge a $\begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}$. Encuentre los vectores característicos de A^{-1}. ¿A qué converge el método de potencias inverso $u_{-k} = A^{-k} u_0$ (después de multiplicar por 0.6^k)?

5. Demuestre que para dos vectores distintos de la misma longitud, $\|x\| = \|y\|$, la transformación de Householder con $v = x - y$ proporciona $Hx = y$ y $Hy = x$.

6. Calcule $\sigma = \|x\|$, $v = x + \sigma z$, y $H = I - 2vv^{\mathrm{T}}/v^{\mathrm{T}}v$. Compruebe $Hx = -\sigma x$:

$$x = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \quad \text{y} \quad z = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

7. Use el problema 6 para encontrar la HAH^{-1} tridiagonal que es semejante a

$$A = \begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 0 \end{bmatrix}.$$

8. Demuestre que empezando con $A_0 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$, el algoritmo QR sin desplazar sólo produce la modesta mejoría $A_1 = \frac{1}{5}\begin{bmatrix} 14 & -3 \\ -3 & 6 \end{bmatrix}$.

9. Aplique a la siguiente matriz A un solo paso QR con el desplazamiento $\alpha = a_{22}$, que en este caso significa sin desplazamiento, ya que $a_{22} = 0$. Demuestre que los elementos fuera de la diagonal van de $\operatorname{sen}\theta$ a $-\operatorname{sen}^3\theta$, que es *convergencia cúbica*.

$$A = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & 0 \end{bmatrix}.$$

10. Compruebe que el algoritmo QR no modifica la matriz tridiagonal $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Es uno de los (raros) contraejemplos de la convergencia (por lo que se desplaza).

11. Demuestre por inducción que, sin desplazamientos, $(Q_0 Q_1 \cdots Q_k)(R_k \cdots R_1 R_0)$ es exactamente la factorización QR de A_{k+1}. Esta identidad relaciona QR con el método de potencias y conduce a una explicación de su convergencia. Si $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$, estos valores característicos aparecen gradualmente sobre la diagonal principal.

12. Escoja $\operatorname{sen}\theta$ y $\cos\theta$ en la rotación P para triangularizar A, y encuentre R:

$$P_{21}A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} = R.$$

13. Escoja $\operatorname{sen}\theta$ y $\cos\theta$ para hacer triangular a $P_{21} A P_{21}^{-1}$ (misma A). ¿Cuáles son los valores característicos?

14. Cuando A se multiplica por P_{ij} (rotación del plano), ¿qué elementos cambian? Cuando P_{ij} se multiplica por la derecha por P_{ij}^{-1}, ¿qué elementos cambian ahora?

15. ¿Cuántas multiplicaciones y cuántas sumas se usan para calcular PA? (Una organización cuidadosa de todas las rotaciones proporciona $\frac{2}{3}n^3$ multiplicaciones y sumas, lo mismo que para QR por reflectores y el doble que para LU.)

16. (Giro de una mano robot) Un robot produce cualquier rotación A de 3 por 3, a partir de rotaciones del plano alrededor de los ejes x, y, y z. Si $P_{32}P_{31}P_{21}A = I$, las tres vueltas del robot están en $A = P_{21}^{-1} P_{31}^{-1} P_{32}^{-1}$. Los tres ángulos son ***ángulos de Euler***. Escoja la primera θ de modo que

$$P_{21}A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \frac{1}{3}\begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix} \text{ es cero en la posición (2, 1)}$$

7.4 MÉTODOS ITERATIVOS PARA $Ax = b$

En contraste con los valores característicos, para los cuales no hay opción, en absoluto se requiere un método iterativo para resolver $Ax = b$. La eliminación gaussiana llega a la solución x en un número finito de pasos ($n^3/3$ para una matriz completa, menos para las grandes matrices que se encuentran en realidad). A menudo este número es razonable. Cuando es enorme, quizá convenga detenerse en una x fija que pueda obtenerse más rápido, y entonces ya no es de utilidad aplicar parcialmente la eliminación y luego detenerse.

Nuestro objetivo es describir métodos que empiecen desde cualquier conjetura inicial x_0, y obtener una aproximación mejorada x_{k+1} a partir de la x_k previa. Es posible detenerse donde se quiera.

Es fácil inventar un método iterativo ***al separar la matriz*** A. Si $A = S - T$, entonces la ecuación $Ax = b$ es la misma que $Sx = Tx + b$. En consecuencia, puede intentarse

$$\textbf{Iteración desde } x_k \textbf{ hasta } x_{k+1} \qquad Sx_{k+1} = Tx_k + b. \tag{1}$$

No hay garantía de que este método funcione. Una separación exitosa $S - T$ satisface dos requisitos:

1. El nuevo vector x_{k+1} debe ser *fácil de calcular*. En consecuencia, S debe ser una matriz sencilla (¡e invertible!); puede ser diagonal o triangular.
2. La sucesión x_k debe *converger* a la solución verdadera x. Si la iteración en la ecuación (1) se resta de la ecuación verdadera $Sx = Tx + b$, el resultado es una fórmula que implica sólo los errores $e_k = x - x_k$:

$$\textbf{Ecuación error} \qquad Se_{k+1} = Te_k. \tag{2}$$

Ésta es justo, una ecuación en diferencias. Empieza con el error inicial e_0, y luego de k pasos produce el nuevo error $e_k = (S^{-1}T)^k e_0$. La cuestión de convergencia es exactamente la misma que la cuestión de estabilidad: $x_k \to x$ exactamente cuando $e_k \to 0$.

7F El método iterativo en la ecuación (1) es ***convergente*** si y sólo si todo valor característico de $S^{-1}T$ satisface $|\lambda| < 1$. Su razón de convergencia depende del tamaño máximo de $|\lambda|$:

$$\textbf{Radio espectral "rho"} \quad \rho(S^{-1}T) = \max_i |\lambda_i|. \tag{3}$$

Recuerde que una solución típica de $e_{k+1} = S^{-1}Te_k$ es una combinación de vectores característicos:

$$\textbf{Error después de } k \textbf{ pasos} \quad e_k = c_1\lambda_1^k x_1 + \cdots + c_n\lambda_n^k x_n. \tag{4}$$

El mayor $|\lambda_i|$ termina por ser dominante, de modo que el radio espectral $\rho = |\lambda_{\text{máx}}|$ gobierna la razón a la que e_k converge a cero. Ciertamente, se requiere $\rho < 1$.

Los requisitos **1** y **2** son contradictorios. Es posible alcanzar convergencia absoluta con $S = A$ y $T = 0$; el primer y único paso de la iteración debe ser $Ax_1 = b$. En ese caso la matriz error $S^{-1}T$ es cero, sus valores característicos y el radio espectral son cero, y la razón de convergencia (que suele definirse como $-\log \rho$) es infinita. Pero $Ax_1 = b$ puede ser difícil de resolver; ésa fue la razón para separar. A menudo una elección simple de S puede ser exitosa, y se empieza con tres posibilidades:

1. S = parte diagonal de A (**método de Jacobi**).
2. S = parte triangular de A (**método de Gauss-Seidel**).
3. S = combinación de 1 y 2 (***sobrerrelajamiento sucesivo*** o SRS).

S también se conoce como ***preacondicionador***, y su elección es crucial en análisis numéricos.

Ejemplo 1 (**Jacobi**) Aquí S es la parte diagonal de A:

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, \quad S = \begin{bmatrix} 2 & \\ & 2 \end{bmatrix}, \quad T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad S^{-1}T = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}.$$

Si las componentes de x son v y w, el paso de Jacobi $Sx_{k+1} = Tx_k + b$ es

$$\begin{aligned} 2v_{k+1} &= w_k + b_1 \\ 2w_{k+1} &= v_k + b_2, \end{aligned} \quad \text{o bien,} \quad \begin{bmatrix} v \\ w \end{bmatrix}_{k+1} = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix}_k + \begin{bmatrix} b_1/2 \\ b_2/2 \end{bmatrix}.$$

La matriz decisiva $S^{-1}T$ tiene valores característicos $\pm\frac{1}{2}$, que significa que el error se reduce a la mitad (un dígito binario más se vuelve correcto) a cada paso. En este ejemplo, que es demasiado pequeño para ser típico, la convergencia es rápida.

Para una matriz A más grande, hay una dificultad muy práctica. **La iteración de Jacobi requiere mantener todas las componentes de x_k hasta que se ha completado el cálculo de x_{k+1}.** Una idea mucho más natural, que sólo requiere la mitad de almacenamiento, es empezar utilizando cada componente de la nueva x_{k+1} tan pronto como es calculada; x_{k+1} toma el lugar de x_k una componente a la vez. Entonces x_k puede destruirse en cuanto x_{k+1} es creada. La primera componente queda como antes:

$$\textbf{Nueva } x_1 \quad a_{11}(x_1)_{k+1} = (-a_{12}x_2 - a_{13}x_3 - \cdots - a_{1n}x_n)_k + b_1.$$

El paso siguiente opera inmediatamente con este nuevo valor de x_1, para encontrar $(x_2)_{k+1}$:

$$\textbf{Nueva } x_2 \quad a_{22}(x_2)_{k+1} = -a_{21}(x_1)_{k+1} + (-a_{23}x_3 - \cdots - a_{2n}x_n)_k + b_2.$$

Y la última ecuación en el paso de iteración usa exclusivamente nuevos valores:

$$\textbf{Nueva } x_n \quad a_{nn}(x_n)_{k+1} = (-a_{n1}x_1 - a_{n2}x_2 - \cdots - a_{nn-1}x_{n-1})_{k+1} + b_n.$$

Esto se denomina ***método de Gauss-Seidel***, aunque aparentemente Gauss no lo conocía y Seidel no lo recomendaba. Esto es parte de una historia sorprendente, ya que no es un mal método. Cuando los términos en x_{k+1} se muevan al miembro izquierdo, *S se ve como la parte triangular inferior de A*. En el miembro derecho, T es estrictamente triangular superior.

Ejemplo 2 (**Gauss-Seidel**) Aquí $S^{-1}T$ tiene valores característicos menores:

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, \quad S = \begin{bmatrix} 2 & 0 \\ -1 & 2 \end{bmatrix}, \quad T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad S^{-1}T = \begin{bmatrix} 0 & \frac{1}{2} \\ 0 & \frac{1}{4} \end{bmatrix}.$$

Un simple paso del método de Gauss-Seidel lleva las componentes v_k y w_k en

$$\begin{aligned} 2v_{k+1} &= w_k + b_1 \\ 2w_{k+1} &= v_{k+1} + b_2, \end{aligned} \quad \text{o bien,} \quad \begin{bmatrix} 2 & 0 \\ -1 & 2 \end{bmatrix} x_{k+1} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x_k + b.$$

Los valores característicos de $S^{-1}T$ son $\frac{1}{4}$ y 0. El error se divide entre 4 cada vez, de modo que *un solo paso de Gauss-Seidel amerita dos pasos de Jacobi*. Debido a que ambos métodos requieren el mismo número de operaciones —simplemente se utiliza el nuevo valor en vez del anterior, y realmente ahorra almacenamiento— el método de Gauss-Seidel es mucho mejor.

Esta regla se cumple en muchas aplicaciones, aun cuando hay ejemplos en los que el método de Jacobi converge y Gauss-Seidel fracasa (o a la inversa). El caso simétrico es directo: cuando todos los $a_{ii} > 0$, Gauss-Seidel converge si y sólo si A es positiva definida.

Durante la época en que los cálculos se hacían manualmente se descubrió (tal vez accidentalmente) que la convergencia es más rápida si se va *más allá* de la corrección $x_{k+1} - x_k$ de Gauss-Seidel. En términos generales, estas aproximaciones quedan en el mismo lado de la solución x. Un ***factor de sobrerrelajamiento ω*** nos acerca más a la solución. Con $\omega = 1$, se recupera Gauss-Seidel; con $\omega > 1$, el método se conoce como ***sobrerrelajamiento sucesivo*** (SRS). La elección óptima de ω nunca es mayor que 2. A menudo está en la vecindad de 1.9.

Para describir el sobrerrelajamiento, sean D, L, y U las partes de A sobre, abajo y arriba de la diagonal, respectivamente. (Esta separación no tiene nada que ver con la $A = LDU$ de la eliminación. De hecho, ahora se tiene $A = L + D + U$.) El método de Jacobi tiene $S = D$ en el miembro izquierdo y $T = -L - U$ en el miembro derecho. El método de Gauss-Seidel escoge $S = D + L$ y $T = -U$. Para acelerar la convergencia, nos movemos a

$$\textbf{Sobrerrelajamiento} \quad [D + \omega L]x_{k+1} = [(1 + \omega)D - \omega U]x_k + \omega b. \tag{5}$$

Sin importar ω, la matriz de la izquierda es triangular inferior y la de la derecha es triangular superior. En consecuencia, x_{k+1} puede seguir sustituyendo a x_k, componente a componente, tan pronto como es calculada. Un paso típico es

$$a_{ii}(x_i)_{k+1} = a_{ii}(x_i)_k + \omega[(-a_{i1}x_1 - \cdots - a_{ii-1}x_{i-1})_{k+1} + (-a_{ii}x_i - \cdots - a_{in}x_n)_k + b_i].$$

Si ocurre que la conjetura anterior x_k coincide con la verdadera solución x, entonces la nueva conjetura x_{k+1} puede ser la misma, y la cantidad entre corchetes se elimina.

Ejemplo 3 (SRS) Para la misma $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$, cada paso de sobrerrelajamiento es

$$\begin{bmatrix} 2 & 0 \\ -\omega & 2 \end{bmatrix} x_{k+1} = \begin{bmatrix} 2(1-\omega) & \omega \\ 0 & 2(1-\omega) \end{bmatrix} x_k + \omega b.$$

Si se divide entre ω, estas dos matrices son S y T en la separación $A = S - T$; la iteración regresa a $Sx_{k+1} = Tx_k + b$. La matriz crucial $L = S^{-1}T$ es

$$L = \begin{bmatrix} 2 & 0 \\ -\omega & 2 \end{bmatrix}^{-1} \begin{bmatrix} 2(1-\omega) & \omega \\ 0 & 2(1-\omega) \end{bmatrix} = \begin{bmatrix} 1-\omega & \frac{1}{2}\omega \\ \frac{1}{2}\omega(1-\omega) & 1-\omega+\frac{1}{4}\omega^2 \end{bmatrix}.$$

La ω óptima hace lo más pequeño posible al mayor valor característico de L (su radio espectral). *Todo el chiste del sobrerrelajamiento es descubrir esta ω óptima.* El producto de los valores característicos es igual a $\det L = \det T/\det S$:

$$\lambda_1\lambda_2 = \det L = (1-\omega)^2.$$

Siempre $\det S = \det D$ porque L está abajo de la diagonal, y $\det T = \det\,(1-\omega)D$ porque U está arriba de la diagonal. Su producto es $\det L = (1-\omega)^n$. (Esto explica por qué nunca se llega tan lejos como a $\omega = 2$. El producto de los valores característicos sería tan grande, y la iteración no convergería.) También se obtiene una pista sobre el comportamiento de los valores característicos: *En la ω óptima, los dos valores característicos son iguales. Ambos deben ser iguales a* $\omega - 1$, *de modo que su producto coincide con* $\det L$. Este valor de ω es fácil de calcular, porque la suma de los valores característicos siempre coincide con la suma de los elementos diagonales (la traza de L):

$$\textbf{ω óptima} \qquad \lambda_1 + \lambda_2 = (\omega_{\text{ópt}} - 1) + (\omega_{\text{ópt}} - 1) = 2 - 2\omega_{\text{ópt}} + \frac{1}{4}\omega_{\text{ópt}}^2. \qquad (6)$$

Esta ecuación cuadrática proporciona $\omega_{\text{ópt}} = 4(2-\sqrt{3}) \approx 1.07$. Los dos valores característicos iguales son aproximadamente $\omega - 1 = 0.07$, que es una reducción importante del valor de Gauss-Seidel $\lambda = \frac{1}{4}$ en $\omega = 1$. En este ejemplo, la elección correcta de ω ha duplicado nuevamente la razón de convergencia, ya que $\left(\frac{1}{4}\right)^2 \approx 0.07$. Si ω se incrementa aún más, los valores característicos se convierten en un par conjugado complejo, ya que ambos tienen $|\lambda| = \omega - 1$, que ahora crece con ω.

El descubrimiento de que tal mejoría podía obtenerse tan fácilmente, casi como por arte de magia, fue el punto de partida para 20 años de intensa actividad en análisis numérico. El primer problema fue resuelto en la tesis de Young de 1950: una simple fórmula para la ω óptima. El paso clave era relacionar los valores característicos λ de L con los valores característicos μ de la matriz original de Jacobi $D^{-1}(-L-U)$. Esta relación se expresa como

$$\textbf{Fórmula para ω} \qquad (\lambda + \omega - 1)^2 = \lambda\omega^2\mu^2. \qquad (7)$$

Lo anterior es válido para una amplia variedad de matrices de diferencias finitas, y si se toma $\omega = 1$ (Gauss-Seidel), se obtiene $\lambda^2 = \lambda\mu^2$. En consecuencia, $\lambda = 0$ y $\lambda = \mu^2$ como en el ejemplo 2, donde $\mu = \pm\frac{1}{2}$ y $\lambda = 0$, $\lambda = \frac{1}{4}$. Todas las matrices en la clase de Young

tienen valores característicos μ que ocurren en pares más-menos, y la λ correspondiente son 0 y μ^2. Así, Gauss-Seidel duplica la razón de convergencia de Jacobi.

El problema importante es escoger ω de modo que $\lambda_{\text{máx}}$ sea minimizado. Por fortuna, ¡la ecuación de Young (7) es exactamente nuestro ejemplo de 2 por 2! La mejor ω hace que ambas raíces λ sean iguales a $\omega - 1$:

$$(\omega - 1) + (\omega - 1) = 2 - 2\omega + \mu^2\omega^2, \quad \text{o bien,} \quad \omega = \frac{2(1-\sqrt{1-\mu^2})}{\mu^2}.$$

Para una matriz grande, este patrón se repite para una cantidad de pares distintos $\pm\mu_i$, y sólo puede hacerse una elección de μ. La μ más grande proporciona el mayor valor de w y de $\lambda = \omega - 1$. Como el objetivo es hacer a $\lambda_{\text{máx}}$ lo más pequeño posible, ese par extremo satisface la mejor elección de $\omega_{\text{ópt}}$:

$$\textbf{ω óptima} \qquad \omega_{\text{ópt}} = \frac{2(1-\sqrt{1-\mu_{\text{máx}}^2})}{\mu_{\text{máx}}^2} \quad \text{y} \quad \lambda_{\text{máx}} = \omega_{\text{ópt}} - 1. \qquad (8)$$

7G La separación de la matriz $-1, 2, -1$ de orden n produce los siguientes valores característicos de B:

Jacobi (matriz $S = 0, 2, 0$): $S^{-1}T$ tiene $|\lambda|_{\text{máx}} = \cos\dfrac{\pi}{n+1}$

Gauss-Seidel (matriz $S = -1, 2, 0$): $S^{-1}T$ tiene $|\lambda|_{\text{máx}} = \left(\cos\dfrac{\pi}{n+1}\right)^2$

SRS (con la mejor ω): $|\lambda|_{\text{máx}} = \left(\cos\dfrac{\pi}{n+1}\right)^2 \Big/ \left(1 + \text{sen}\dfrac{\pi}{n+1}\right)^2.$

Lo anterior sólo puede apreciarse con un ejemplo. Suponga que A es de orden 21, lo cual es bastante moderado. Así, $h = \frac{1}{22}$, $\cos\pi h = 0.99$ y el método de Jacobi es lento; $\cos^2\pi h = 0.98$ significa que incluso el método de Gauss-Seidel requiere muchas iteraciones. Pero como $\text{sen}\,\pi h = \sqrt{0.02} = 0.14$, el método de sobrerrelajamiento óptimo tendrá el factor de convergencia

$$\lambda_{\text{máx}} = \frac{0.86}{1.14} = 0.75, \quad \text{con} \quad \omega_{\text{ópt}} = 1 + \lambda_{\text{máx}} = 1.75.$$

El error se reduce en 25% en cada paso, y ***un solo paso SRS es el equivalente de 30 pasos de Jacobi***: $(0.99)^{30} = 0.75$.

Este es un sorprendente resultado de una idea tan sencilla. Sus aplicaciones reales no son en un problema unidimensional como $-u_{xx} = f$. Un sistema tridiagonal $Ax = b$ ya es fácil. Para lo que es importante el sobrerrelajamiento (y otros conceptos) es para las ecuaciones diferenciales. Al cambiar a $-u_{xx} - u_{yy} = f$ se llega al "esquema de cinco puntos". Los elementos $-1, 2, -1$ en la dirección x se combinan con $-1, 2, -1$ en la dirección y para proporcionar una diagonal principal de $+4$ y cuatro elementos de -1 fuera de la diagonal. ¡*La matriz A no tiene un ancho de banda pequeño*! No hay ninguna forma para numerar los N^2 puntos de la malla en un cuadrado, de modo que cada punto permanezca próximo a sus cuatro vecinos. Esta es la verdadera ***dirección de la dimensionalidad***, y computadoras en paralelo la alivian parcialmente.

Si el ordenamiento se realiza renglón por renglón, cada punto debe esperar todo un renglón para que el vecino de arriba se aproxime. El ancho de banda de la "matriz de cinco puntos" es N:

$$\begin{matrix} -1, 2, -1 \text{ en } x \text{ y } y \\ \text{da } -1, -1, 4, -1, -1 \end{matrix} \qquad A = \left[\quad N \quad \right].$$

Esta matriz ha recibido más atención, y ha sido atacada en más formas diferentes que cualquier otra ecuación lineal $Ax = b$. La tendencia actual es volver a métodos directos, con base en una idea de Golub y Hockney: ciertas matrices especiales se apartarán cuando se aborden de la manera idónea. (Esto es comparable a la Transformada de Fourier Rápida.) Antes del surgimiento de los métodos iterativos de ***dirección alternada***, donde la separación dividía a la matriz tridiagonal en la dirección x y en la matriz en la dirección y. Una elección reciente es $S = L_0U_0$, donde elementos pequeños de las verdaderas L y U se igualan a cero mientras se factoriza A. Se denomina ***LU incompleta*** y puede ser terrible.

No es posible terminar esta sección sin mencionar el ***método del gradiente conjugado***, que parecía muerto aunque repentinamente resultó bastante vivo (los pasos se proporcionan en el problema 33). Es directo, más que iterativo, pero a diferencia de la eliminación, puede detenerse en cualquier parte del proceso. Y aunque no merece la pena mencionarlo, un concepto completamente nuevo puede surgir y ganar. Sin embargo, parece justo decir que fue el cambio de 0.99 a 0.75 lo que revolucionó la solución de $Ax = b$.

Conjunto de problemas 7.4

1. Los valores característicos de la siguiente matriz son $2-\sqrt{2}, 2,$ y $2+\sqrt{2}$:

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}.$$

Encuentre la matriz de Jacobi $D^{-1}(-L-U)$ y la matriz de Gauss-Seidel $(D+L)^{-1}(-U)$ y sus valores característicos, así como los números $\omega_{\text{ópt}}$ y $\lambda_{\text{máx}}$ para SRS.

2. Para la siguiente matriz de n por n, describa la matriz de Jacobi $J = D^{-1}(-L-U)$:

$$A = \begin{bmatrix} 2 & -1 & & \\ -1 & \cdot & \cdot & \\ & \cdot & \cdot & -1 \\ & & -1 & 2 \end{bmatrix}.$$

Demuestre que el vector $x_1 = (\text{sen } \pi h, \text{sen } 2\pi h, \ldots, \text{sen } n\pi h)$ es un vector característico de J con valor característico $\lambda_1 = \cos \pi h = \cos \pi/(n+1)$.

3. En el problema 2, demuestre que $x_k = (\text{sen } k\pi h, \text{sen } 2k\pi h, \ldots, \text{sen } nk\pi h)$ es un vector característico de A. Multiplique x_k por A para encontrar el valor característico correspondiente α_k. Compruebe que en el caso de 3 por 3 estos valores característicos son $2-\sqrt{2}, 2, 2+\sqrt{2}$.

Nota Los valores característicos de la matriz de Jacobi $J = \frac{1}{2}(-L-U) = I - \frac{1}{2}A$ son $\lambda_k = 1 - \frac{1}{2}\alpha_k = \cos k\pi h$. Ocurren en pares más-menos y $\lambda_{\text{máx}}$ es $\cos \pi h$.)

Los problemas 4 y 5 requieren el "teorema del círculo" de Gershgorin: *Todo valor característico de A está en por lo menos uno de los círculos $C_1, \ldots, C_n$, donde C_i tiene su centro en el elemento diagonal a_{ii}. Su radio $r_i = \sum_{j \neq i} |a_{ij}|$ es igual a la suma absoluta a lo largo del resto del renglón.*

Demostración Suponga que x_i es la mayor componente de x. Entonces $Ax = \lambda x$ conduce a

$$(\lambda - a_{ii})x_i = \sum_{j \neq i} a_{ij}x_j, \quad \text{o bien,} \quad |\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}| \frac{|x_j|}{|x_i|} \leq \sum_{j \neq i} |a_{ij}| = r_i. \quad \blacksquare$$

4. La matriz

$$A = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 4 & 1 \\ 2 & 2 & 5 \end{bmatrix}$$

Se denomina *dominante diagonalmente* porque toda $|a_{ii}| > r_i$. Demuestre que cero no puede estar en ninguno de los círculos, y concluya que A es no singular.

5. Escriba la matriz de Jacobi J para la matriz A diagonalmente dominante del problema 4, y encuentre los tres círculos de Gershgorin para J. Demuestre que todos los radios satisfacen $r_i < 1$, y que la iteración de Jacobi converge.

6. La verdadera solución de $Ax = b$ es ligeramente diferente de la solución por eliminación de $LUx_0 = b$; $A - LU$ se pierde el cero debido al redondeo. Una estrategia es hacer todo en doble precisión, aunque una manera mejor y más rápida es el *refinamiento iterativo*: Calcula sólo un vector $r = b - Ax_0$ en doble precisión, resuelva $LUy = r$, y sume la corrección y a x_0. Problema: Multiplicar $x_1 = x_0 + y$ por LU, escriba el resultado como una separación $Sx_1 = Tx_0 + b$, y explique por qué T es extremadamente pequeña. Este simple paso lleva casi exactamente a x.

7. Para una matriz general de 2 por 2

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

encuentre la matriz de la iteración de Jacobi $S^{-1}T = -D^{-1}(L+U)$ y sus valores característicos μ_i. También encuentre la matriz de Gauss-Seidel $-(D+L)^{-1}U$ y sus valores característicos λ_i, y decida si $\lambda_{\text{máx}} = \mu^2_{\text{máx}}$.

8. Cambie $Ax = b$ a $x = (I - A)x + b$. ¿Cuáles son S y T para esta separación? ¿Qué matriz $S^{-1}T$ controla la convergencia de $x_{k+1} = (I-A)x_k + b$?

9. Si λ es un valor característico de A, entonces ______ es un valor característico de $B = I - A$. Los verdaderos valores característicos de B tienen valor absoluto menor que 1 si los verdaderos valores característicos de A están entre ______ y ______.

10. Demuestre por qué la iteración $x_{k+1} = (I-A)x_k + b$ no converge para $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$.

11. ¿Por qué la norma de B^k nunca es mayor que $\|B\|^k$? Entonces $\|B\| < 1$ garantiza que las potencias de B^k tienden a cero (convergencia). Esto no es una sorpresa, ya que $|\lambda|_{\text{máx}}$ está debajo de $\|B\|$.

12. Si A es singular, entonces deben fracasar todas las separaciones $A = S - T$. A partir de $Ax = 0$, demuestre que $S^{-1}Tx = x$. Por tanto, esta matriz $B = S^{-1}T$ tiene $\lambda = 1$ y fracasa.

13. Cambie los 2s a 3s, y encuentre los valores característicos de $S^{-1}T$ por ambos métodos:

$$\textbf{(J)} \quad \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} x_{k+1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x_k + b \qquad \textbf{(GS)} \quad \begin{bmatrix} 3 & 0 \\ -1 & 3 \end{bmatrix} x_{k+1} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x_k + b.$$

¿$|\lambda|_{\text{máx}}$ para Gauss-Seidel es igual a $|\lambda|^2_{\text{máx}}$ para Jacobi?

14. Escriba un código de computadora (MATLAB u otro) para Gauss-Seidel. Puede definir S y T a partir de A, o establecer el ciclo iterativo directamente a partir de los elementos a_{ij}. Pruébelo en $-1, 2, -1$ las matrices A de orden 10, 20, 50, con $b = (1, 0, \ldots, 0)$.

15. La matriz de separación S para **SRS** es la misma que para Gauss-Seidel, excepto que la diagonal se dividió entre ω. Escriba un programa para **SRS** de una matriz de n por n. Aplíquelo con $\omega = 1, 1.4, 1.8, 2.2$ cuando A es la matriz $-1, 2, -1$ de orden 10.

16. Cuando $A = A^{\mathrm{T}}$, el *método de Arnoldi-Lanczos* encuentra qs ortonormales de modo que $Aq_j = b_{j-1}q_{j-1} + a_jq_j + b_jq_{j+1}$ (con $q_0 = 0$). Multiplique por q_j^T con la finalidad de encontrar una fórmula para a_j. La ecuación indica que $AQ = QT$, donde T es una matriz ________.

17. ¿Cuál es la cota sobre $|\lambda|_{\text{máx}}$ proporcionada por Gershgorin para las siguientes matrices (consulte el problema 4)? ¿Cuáles son los tres círculos de Gershgorin que contienen a todos los valores característicos?

$$A = \begin{bmatrix} 0.3 & 0.3 & 0.2 \\ 0.3 & 0.2 & 0.4 \\ 0.2 & 0.4 & 0.1 \end{bmatrix} \qquad A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}.$$

La cuestión clave para matrices grandes, es que la multiplicación matriz-vector es mucho más rápida que la multiplicación matriz-matriz. Una construcción crucial empieza con un vector b y calcula $Ab, A^2b, \ldots$ (¡pero nunca A^2!) Los N primeros vectores generan el N-ésimo ***subespacio de Krilov***. Se trata de las columnas de la ***matriz K_N de Krilov***:

$$K_N = \begin{bmatrix} b & Ab & A^2b & \cdots & A^{N-1}b \end{bmatrix}.$$

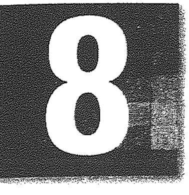
La iteración de Arnoldi-Lanczos ortogonaliza las columnas de K_N, y la iteración del gradiente conjugado resuelve $Ax = b$ cuando A es positiva definida simétrica.

Iteración de Arnoldi	**Iteración del gradiente conjugado**	
$q_1 = b/\lVert b\rVert$	$x_0 = 0, r_0 = b, p_0 = r_0$	
para $n = 1$ a $N - 1$	para $n = 1$ a N	
$v = Aq_n$	$\alpha_n = (r_{n-1}^{\mathrm{T}}r_{n-1})/(p_{n-1}^{\mathrm{T}}Ap_{n-1})$	longitud del paso x_{n-1} a x_n
para $j = 1$ a n	$x_n = x_{n-1} + \alpha_n p_{n-1}$	aproxime la solución
$h_{jn} = q_j^{\mathrm{T}}v$	$r_n = r_{n-1} - \alpha_n Ap_{n-1}$	residuo nuevo $b - Ax_n$
$v = v - h_{jn}q_j$	$\beta_n = (r_n^{\mathrm{T}}r_n)/(r_{n-1}^{\mathrm{T}}r_{n-1})$	mejora en este paso
$h_{n+1,n} = \lVert v\rVert$	$p_n = r_n + \beta_n p_{n-1}$	dirección de la siguiente búsqueda
$q_{n+1} = v/h_{n+1,n}$	*Nota: Sólo 1 multiplicación matriz-vector Aq y Ap.*	

18. En Arnoldi, demuestre que q_2 es ortogonal a q_1. El método de Arnoldi es la ortogonalización de Gram-Schmidt aplicada a la matriz de Krilov: $K_N = Q_NR_N$. Los valores característicos de $Q_N^{\mathrm{T}}AQ_N$ a menudo se encuentran muy próximos a los de A, incluso para $N \ll n$. La *iteración de Lanczos* es el método de Arnoldi para matrices simétricas (todas codificadas en ARPACK).

19. En gradientes conjugados, demuestre que r_1 es ortogonal a r_0 (residuos ortogonales), y $p_1^{\mathrm{T}}Ap_0 = 0$ (las direcciones de búsqueda son A-ortogonales). La iteración resuelve $Ax = b$ al minimizar el error $e^{\mathrm{T}}Ae$ en el subespacio de Krilov. Es un algoritmo fantástico.

Capítulo



Programación lineal y teoría de juegos

8.1 DESIGUALDADES LINEALES

El álgebra concierne ecuaciones, y a menudo el análisis concierne desigualdades. La línea entre ambos siempre ha parecido clara. Sin embargo, me he percatado que este capítulo es un contraejemplo: ***la programación lineal es sobre desigualdades***, aunque constituye una parte incuestionable del álgebra lineal. También es extremadamente útil: es más probable que las decisiones de negocios impliquen programación lineal que determinantes o valores característicos.

Hay tres formas para abordar las matemáticas subyacentes: intuitivamente a través de la geometría, computacionalmente a través del método simplex, o algebraicamente a través de la dualidad. Estos enfoques se desarrollan en las secciones 8.1, 8.2, y 8.3. Luego, la sección 8.4 es sobre problemas (como el matrimonio) donde la solución es un entero. En la sección 8.5 se analizan el póquer y otros juegos matriciales. Los estudiantes del MIT en *Bringing Down the House* contaron cartas altas para ganar en blackjack (en Las Vegas se siguen reglas fijas, y un verdadero juego matricial implica estrategias al azar).

En la sección 8.3 se presenta algo nuevo en esta cuarta edición. El método simplex ahora se presenta en franca competencia con una forma completamente nueva de hacer los cálculos, denominada **método del punto interior**. La emoción comenzó cuando Karmarkar afirmó que su versión era 50 veces más rápida que el método simplex. (Su algoritmo, delineado en la sección 8.2, fue uno de los primeros que fueron patentados: algo que entonces creíamos imposible, y no realmente deseable.) Esa afirmación provocó una oleada de investigación sobre métodos que tienden a la solución desde el "interior", donde todas las desigualdades son estrictas: $x \geq 0$ se convierte en $x > 0$. El resultado es ahora una gran forma de obtener ayuda del problema dual para resolver el problema primario.

Una clave para este capítulo es considerar los significados geométricos de las ***desigualdades lineales***. Una desigualdad divide el espacio n-dimensional en un ***semiespacio*** donde se cumple la desigualdad, y en un subespacio en el que no lo hace. Un ejemplo típico es $x + 2y \geq 4$. La frontera entre los dos semiespacios es la recta $x + 2y = 4$, donde la desigualdad es "apretada". La figura 8.1 se vería casi igual en tres dimensiones. La frontera se convierte en un plano como $x + 2y + z = 4$, y arriba se encuentra el semiespacio $x + 2y + z \geq 4$. En n dimensiones, la "dimensión" del plano es $n - 1$.

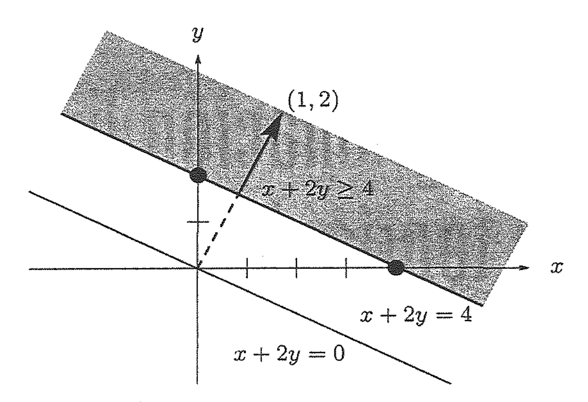


Figura 8.1 Las ecuaciones proporcionan rectas y planos. Las desigualdades, semiespacios.

Hay otra restricción fundamental para la programación lineal: se requiere que x y y sean ***nonegativas***. Este par de desigualdades $x \geq 0$ y $y \geq 0$ produce dos subespacios más. La figura 8.2 está acotada por los ejes coordenados: $x \geq 0$ admite todos los puntos a la derecha de $x = 0$, y $y \geq 0$ es el subespacio que está arriba de $y = 0$.

El conjunto factible y la función de costo

El paso importante es imponer de una vez las tres desigualdades. Éstas se combinan para proporcionar la región sombreada de la figura 8.2. El ***conjunto factible*** es la *intersección* de los tres subespacios $x + 2y \geq 4$, $x \geq 0$, y $y \geq 0$. Un conjunto factible está constituido por las soluciones de una familia de desigualdades lineales como $Ax \geq b$ (la intersección de m subespacios). Cuando también se requiere que toda componente de x sea no negativa (la desigualdad vectorial $x \geq 0$), se agregan n subespacios más. Mientras más restricciones se imponen, más pequeño es el espacio factible.

Puede ocurrir fácilmente que un conjunto esté acotado o incluso sea vacío. Si el ejemplo se cambia al subespacio $x + 2y \leq 4$, preservando $x \geq 0$ y $y \geq 0$, se obtiene el pequeño triángulo OAB. Al combinar las dos desigualdades $x + 2y \geq 4$ y $x + 2y \leq 4$, el conjunto se reduce a una recta donde $x + 2y = 4$. Si se añade una restricción contradictoria como $x + 2y \leq -2$, el conjunto factible es vacío.

El álgebra de las desigualdades lineales (o conjuntos factibles) constituye una parte de nuestro tema de estudio. Sin embargo, la programación lineal tiene otros ingredientes: busca *el punto factible que maximiza o minimiza una cierta* ***función de costo*** como $2x + 3y$. El problema en programación lineal es encontrar el punto que ***está en el conjunto factible y minimiza el costo.***

El problema se ilustra con la geometría de la figura 8.2. La familia de costos $2x + 3y$ proporciona una familia de rectas paralelas. El costo mínimo aparece cuando la primera recta corta al conjunto factible. Esta intersección ocurre en B, donde $x^* = 0$ y $y^* = 2$; el costo mínimo es $2x^* + 3y^* = 6$. El vector $(0, 2)$ es ***factible*** porque está en el conjunto factible, y es ***óptimo*** porque minimiza la función de costo, y el costo mínimo 6 es el ***valor*** del programa. Los vectores óptimos se denotan con un asterisco.

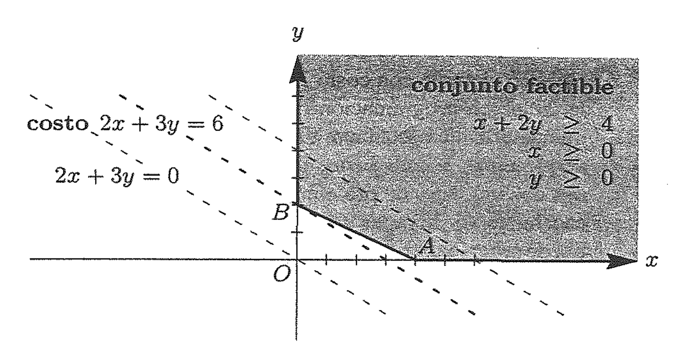


Figura 8.2 Conjunto factible con lados planos, y los costos $2x + 3y$, que se tocan en B.

El vector óptimo ocurre en un vértice del conjunto factible. Este hecho lo garantiza la geometría, ya que las rectas que proporcionan la función de costo (o los planos, cuando se llega a más incógnitas) se desplazan de manera continua hasta que cortan al conjunto factible. ¡El primer contacto debe ocurrir a lo largo de su frontera! El "método simplex" va de un vértice del conjunto factible al siguiente hasta que encuentra el vértice de menor costo. En contraste, los "métodos del punto interior" tienden a esa solución óptima desde ***el interior*** del conjunto factible.

Nota Con una función de costo distinta, la intersección podría no ser un solo punto. Si el costo fuese $x + 2y$, toda la arista entre B y A sería óptima. El costo mínimo es $x^* + 2y^*$, que es igual a 4 para todos estos vectores óptimos. En nuestro conjunto factible, ¡el problema máximo no tendría solución! El costo podría elevarse arbitrariamente y el costo máximo sería infinito.

Todo problema de programación lineal puede ubicarse en cualquiera de las tres categorías posibles siguientes:

1. El conjunto factible es *vacío.*
2. La función de costo *no está acotada* en el conjunto factible.
3. El costo alcanza su *mínimo* (o máximo) en el conjunto factible: *este es el caso bondadoso.*

Los casos vacío y no acotado son bastante raros para un problema genuino en economía o ingeniería. Se espera una solución.

Variables flojas

Hay una forma simple para cambiar la desigualdad $x + 2y \geq 4$ en una ecuación. Simplemente se introduce la diferencia como una ***variable floja*** $w = x + 2y - 4$. ¡Esta es nuestra ecuación! La restricción anterior $x + 2y \geq 4$ se convierte en $w \geq 0$, lo cual coincide perfectamente con las otras restricciones de la desigualdad $x \geq 0$, $y \geq 0$. Así, se tienen sólo ecuaciones y simples restricciones de no negatividad sobre x, y, w. Las variables w que "quitan la flojera" ahora se incluyen en el vector incógnito x:

Problema original ***Minimizar*** cx ***sujeto a*** $Ax = b$ ***y*** $x \geq 0$.

El vector renglón c contiene los costos; en nuestro ejemplo, $c = [2\ 3\ 0]$. La condición $x \geq 0$ pone el problema en la parte no negativa de $\mathbf{R}^n$. Estas desigualdades se reducen a las soluciones de $Ax = b$. La eliminación corre peligro, por lo que se requiere una idea completamente nueva.

El problema de la dieta y su dual

Nuestro ejemplo con costo $2x + 3y$ puede plantearse con palabras. Ilustra el "problema de la dieta" en programación lineal, con dos fuentes de proteínas, por ejemplo, un bistec y mantequilla de cacahuate. Cada libra de mantequilla de cacahuate proporciona una unidad de proteínas, y cada bistec proporciona dos unidades. En la dieta se requieren por lo menos dos unidades de proteínas. Por tanto, una dieta que contiene x libras de mantequilla de cacahuate y y bisteces está restringida por $x + 2y \geq 4$, así como por $x \geq 0$ y $y \geq 0$. (No puede tenerse un bistec o mantequilla de cacahuate negativos.) Este es el conjunto factible, y el problema es minimizar el costo. Si una libra de mantequilla de cacahuate cuesta \$2 y un bistec cuesta \$3, entonces el costo de toda la dieta es $2x + 3y$. Por fortuna, la dieta óptima consta de dos bisteces: $x^* = 0$ y $y^* = 2$.

Todo programa lineal, incluyendo éste, tiene un ***dual***. Si el problema original es una minimización, su dual es una maximización. ***El mínimo en el "problema original" dado, es igual al máximo en su dual.*** Esta es la clave de la programación lineal, que se explicará en la sección 8.3. Aquí se permanece con el problema de la dieta y se intenta interpretar su dual.

En lugar del comprador, quien compra suficientes proteínas al costo mínimo, el problema dual es confrontado por el farmacéutico. Las *píldoras de proteínas* compiten con los bisteces y con la mantequilla de cacahuate. De inmediato se encuentran los dos ingredientes de un programa lineal típico: el farmacéutico maximiza el precio p de las píldoras, aunque este precio está sujeto a restricciones lineales. Las proteínas sintéticas no deben costar más que las proteínas que hay en la mantequilla de cacahuate (\$2 por unidad) o las proteínas que hay en los bisteces (\$3 por dos unidades). El precio debe ser no negativo o el farmacéutico no venderá nada. Debido a que se requieren cuatro unidades de proteínas, la ganancia del farmacéutico es de $4p$:

Problema dual ***Maximizar $4p$, sujeto a $p \leq 2$, $2p \leq 3$ y $p \geq 0$.***

En este ejemplo es más fácil resolver el dual que el original; sólo tiene una incógnita p. La restricción $2p \leq 3$ es la estricta que en realidad es activa, y el precio máximo de la proteína sintética es $p = \$1.50$. La ganancia máxima es $4p = \$6$, y el comprador termina pagando lo mismo por la proteína natural que por la proteína sintética. Este es el teorema de dualidad: ***el máximo es igual al mínimo***.

Aplicaciones típicas

La siguiente sección se concentra en la resolución de programas lineales. Este es el momento para describir dos situaciones prácticas en las que se *minimiza o maximiza una función de costo lineal sujeta a restricciones lineales*.

1. Planeación de la producción. Suponga que la General Motors gana \$200 por cada Chevrolet, \$300 por cada Buick y \$500 por cada Cadillac. Estos automóviles dan 20, 17, y 14 millas por galón, respectivamente, y el Congreso insiste en que el automóvil promedio debe proporcionar 18. La planta puede ensamblar un Chevrolet en un minuto, un Buick en dos minutos, y un Cadillac en 3 minutos. ¿Cuál es la ganancia máxima en 8 horas (480 minutos)?

Problema *Maximizar la ganancia* $200x + 300y + 500z$ *sujeta a*

$$20x + 17y + 14z \geq 18(x + y + z), \quad x + 2y + 3z \leq 480, \quad x, y, z \geq 0.$$

2. Selección del portafolios. Los bonos federales pagan 5%; los municipales, 6%, y los bonos chatarra, 9%. Es posible comprar cantidades x, y, z sin exceder un total de \$100 000. El problema consiste en maximizar el interés, con dos restricciones:

i) no es posible invertir más de \$20 000 en bonos chatarra, y
ii) la calidad media del portafolios no debe ser menor que los bonos municipales, de modo que $x \geq z$.

Problema *Maximizar* $5x + 6y + 9z$ *sujeto a*

$$x + y + z \leq 100000, \quad z \leq 20000, \quad z \leq x, \quad x, y, z \geq 0.$$

Las tres desigualdades proporcionan tres variables flojas, con nuevas ecuaciones como $w = x - z$ y desigualdades $w \geq 0$.

Conjunto de problemas 8.1

1. Trace el conjunto factible con restricciones $x + 2y \geq 6$, $2x + y \geq 6$, $x \geq 0$, $y \geq 0$. ¿Qué puntos están en los tres "vértices" de este conjunto?

2. (Recomendado) Sobre el conjunto factible precedente, ¿cuál es el valor mínimo de la función de costo $x + y$? Trace la recta $x + y =$ constante que toca primero el conjunto factible. ¿Qué puntos minimizan las funciones de costo $3x + y$ y $x - y$?

3. Demuestre que el conjunto factible restringido por $2x + 5y \leq 3$, $-3x + 8y \leq -5$, $x \geq 0$, $y \geq 0$, es vacío.

4. Demuestre que el siguiente problema es factible pero no acotado, de modo que no tiene solución óptima: Maximizar $x + y$, sujeto a $x \geq 0$, $y \geq 0$, $-3x + 2y \leq -1$, $x - y \leq 2$.

5. Agregue una simple desigualdad restrictiva a $x \geq 0$, $y \geq 0$ de modo que el conjunto factible contenga sólo un punto.

6. ¿Qué forma tiene el conjunto factible $x \geq 0$, $y \geq 0$, $z \geq 0$, $x + y + z = 1$, y cuál es el máximo de $x + 2y + 3z$?

7. Resuelva el problema del portafolios al final de la sección precedente.

8. En el conjunto factible para el problema de la General Motors, la no negatividad $x, y, z \geq 0$ deja un octavo del espacio tridimensional (el octante positivo). ¿Cómo es cortado esto por los dos planos de las restricciones, y qué forma tiene el conjunto factible? ¿Cómo muestran sus vértices que, con sólo estas dos restricciones, sólo hay dos tipos de automóviles en la solución óptima?

9. (Problema de transporte) Suponga que Texas, California, y Alaska producen —cada uno— un millón de barriles de petróleo; en Chicago se requieren 800 000 barriles, a una distancia de 1000, 2000, y 3000 millas de los tres productores, respectivamente; y se necesitan 2 200 000 barriles en Nueva Inglaterra a 1500, 3000, y 3700 millas de distancia. Si el embarque cuesta una unidad por cada barril-milla, ¿qué programa lineal con cinco restricciones de igualdad debe resolverse con la finalidad de minimizar el costo por embarque?

8.2 EL MÉTODO SIMPLEX

Esta sección aborda la programación lineal con n incógnitas $x \geq 0$ y m restricciones $Ax \geq b$. En la sección previa se tenían dos variables, y una restricción $x + 2y \geq 4$. No es difícil explicar todo el problema, pero no es fácil resolverlo.

El mejor método es plantear el problema en forma matricial. Se cuenta con A, b, y c:

1. una matriz A de m por n,
2. un vector columna b con m componentes, y
3. un vector renglón c (**vector de costo**) con n componentes.

Para ser "factible", el vector x debe satisfacer $x \geq 0$ y $Ax \geq b$. El vector óptimo x^* es el *vector factible de costo mínimo*, y el costo es $cx = c_1x_1 + \cdots + c_nx_n$.

Problema mínimo *Minimizar el costo cx, sujeto a $x \geq 0$ y $Ax \geq b$.*

La condición $x \geq 0$ restringe a x al cuadrante positivo en el espacio n-dimensional. En $\mathbf{R}^2$ es un cuarto del plano; es un octavo de $\mathbf{R}^3$. Un vector aleatorio tiene una posibilidad en 2^n de ser no negativo. $Ax \geq b$ produce m semiespacios adicionales, y los vectores factibles satisfacen todas las $m + n$ condiciones. En otras palabras, x está en la intersección de $m + n$ semiespacios. Este **conjunto factible** tiene lados planos; puede no estar acotado y ser vacío.

La función de costo cx aporta al problema una familia de planos paralelos. Un plano $cx = 0$ pasa por el origen. Los planos $cx =$ constante proporcionan todos los costos posibles. A medida que varía el costo, estos planos barren todo el espacio n-dimensional. *La x^* óptima (el menor costo) ocurre en el punto en que **los planos tocan por vez primera al conjunto factible**.*

Nuestro objetivo es calcular x^*. Esto puede hacerse (en principio) encontrando todos los vértices del conjunto factible, y calculando sus costos. En la práctica esto es imposible. Podría haber millones de vértices, por lo que sería imposible calcularlos todos. En vez de hacer lo anterior, se aplica el ***método simplex***, que es uno de los conceptos más célebres en matemáticas computacionales. Este método fue desarrollado por Dantzig como una forma sistemática para resolver programas lineales, y ya sea por suerte o no, se trata de un éxito sorprendente. Los pasos del método simplex se resumirán en breve, ya que primero se intentará explicarlos.

La geometría: movimiento a lo largo de las aristas

Considero que lo que revela el método es la explicación geométrica. En la fase I simplemente se localiza un vértice del conjunto factible. ***El meollo del método va de vértice a vértice a lo largo de las aristas del conjunto factible.*** En un vértice típico hay n aristas a escoger. Algunas aristas se alejan de la x^* óptima pero desconocida, y otras llevan gradualmente hacia ésta. Dantzig escogió una arista que lleva a un nuevo vértice con un *costo menor*. No hay posibilidad de volver a nada que sea más costoso. Finalmente se llega a un vértice especial, a partir del cual todas las aristas conducen al camino equivocado: se ha minimizado el costo. Este vértice es el vector óptimo x^*, y ahí se detiene el método.

El siguiente problema es transformar los conceptos de *vértice* y *arista* en álgebra lineal. **Una arista es el punto de encuentro de n planos distintos.** Cada plano está dado por una ecuación —justo como tres planos (muro frontal, muro lateral, y piso) producen un vértice en tres dimensiones. Cada vértice del plano factible proviene de transformar n de las $n + m$ desigualdades $Ax \geq b$ y $x \geq 0$ en ecuaciones, y encontrar la intersección de estos n planos.

Una posibilidad es escoger las n ecuaciones $x_1 = 0, \ldots, x_n = 0$, y terminar en el origen. Así como todas las demás posibilidades, *este punto de intersección sólo es un vértice genuino si también satisface las demás m restricciones de desigualdad.* En caso contrario, ni siquiera está en el conjunto factible, por lo que es un engaño total. Nuestro ejemplo con $n = 2$ variables y $m = 2$ restricciones tiene seis intersecciones, que se ilustran en la figura 8.3. Tres de ellas en realidad son los vértices P, Q, R del conjunto factible. Se trata de los vectores (0, 6), (2, 2) y (6, 0). Uno de ellos debe ser el vector óptimo (a menos que el costo mínimo sea $-\infty$). Los otros tres, incluyendo el origen, son falsos.

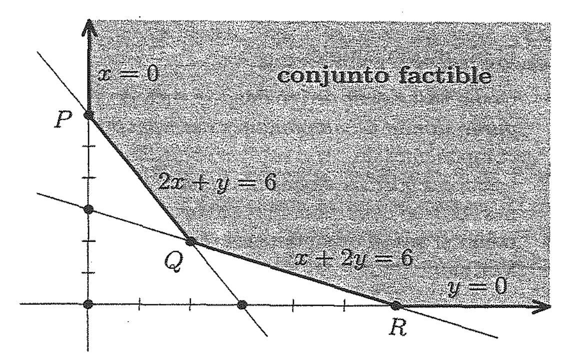


Figura 8.3 Los vértices P, Q, R y las aristas del conjunto factible.

En general, hay $(n + m)!/n!m!$ intersecciones posibles. Esto incluye el número de formas para elegir n ecuaciones planas de $n + m$. El tamaño de este coeficiente binomial hace imposible el cálculo de todos los vértices para m y n grandes. La tarea de la Fase I es encontrar un vértice genuino o establecer que el conjunto factible es vacío. Continuamos con la hipótesis de que se ha encontrado un vértice.

Suponga que se retira uno de los n planos de intersección. **Los puntos que satisfacen las $n - 1$ ecuaciones restantes constituyen una arista que sale del vértice.** Este vértice es la intersección de los $n - 1$ planos. Para permanecer en el conjunto factible, a lo largo de cada arista sólo se permite una dirección. Pero se tiene una elección de n aristas, y la Fase II debe hacer esta elección.

Para describir esta fase, $Ax \geq b$ vuelve a escribirse en una forma completamente paralela a las n restricciones simples $x_j \geq 0$. Este es el papel de las ***variables flojas*** $w = Ax - b$. Las restricciones $Ax \geq b$ se traducen a $w_1 \geq 0, \ldots, w_m \geq 0$, con una variable floja para cada renglón de A. La ecuación $w = Ax - b$, o $Ax - w = b$, pasa a la forma matricial:

Las variables flojas proporcionan m ecuaciones $\quad \begin{bmatrix} A & -I \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} = b.$

El conjunto factible está regido por estas m ecuaciones y las $n + m$ desigualdades simples $x \geq 0$, $w \geq 0$. Ahora se cuenta con ***restricciones de igualdad y nonegatividad***.

El método simplex no establece ninguna diferencia entre x y w, de modo que se simplifica:

$\begin{bmatrix} A & -I \end{bmatrix}$ se vuelve a nombrar como A $\quad \begin{bmatrix} x \\ w \end{bmatrix}$ se vuelve a nombrar como x $\quad \begin{bmatrix} c & 0 \end{bmatrix}$ se vuelve a nombrar como c.

Las restricciones de igualdad son ahora $Ax = b$. Las $n + m$ desigualdades se vuelven justo $x \geq 0$. El único rastro que queda de la variable floja w se encuentra en el hecho de que la nueva matriz A es de m por $n + m$, y la nueva x tiene $n + m$ componentes. Se preserva mucho de la notación original, dejando sin cambio a m y n como recordatorio de lo que ha ocurrido. El problema ha cambiado a: ***Minimizar cx, sujeto a $x \geq 0$ y $Ax = b$.***

Ejemplo 1 El problema en la figura 8.3 tiene las restricciones $x + 2y \geq 6$, $2x + y \geq 6$, y costo $x + y$. El nuevo sistema tiene cuatro incógnitas (x, y, y dos variables flojas):

$$A = \begin{bmatrix} 1 & 2 & -1 & 0 \\ 2 & 1 & 0 & -1 \end{bmatrix} \qquad b = \begin{bmatrix} 6 \\ 6 \end{bmatrix} \qquad c = \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix}.$$

El algoritmo simplex

Con restricciones de igualdad, el método simplex puede comenzar. ***Un vértice es ahora un punto donde n componentes del nuevo vector x*** (los x y w anteriores) ***son cero.*** Estas n componentes de x son las *variables libres* en $Ax = b$. Las m componentes restantes son las *variables básicas* o *variables pivote*. Al igualar a cero las n variables libres, las m ecuaciones $Ax = b$ determinan las m variables básicas. Esta "solución básica" x es un vértice genuino si sus m componentes distintas de cero son *positivas*. Así, x pertenece al conjunto factible.

8A Los ***vértices del conjunto factible*** son las ***soluciones factibles básicas*** de $Ax = b$. Una solución es *básica* cuando n de sus $m + n$ componentes son cero, y es *factible* cuando satisface $x \geq 0$. La fase I del método simplex encuentra una solución factible básica. La fase II se desplaza paso a paso hacia la x^* óptima.

El punto vértice P en la figura 8.3 es la intersección de $x = 0$ con $2x + y - 6 = 0$.

Vértice **(0, 6, 6, 0)**
Básica **(dos ceros)**
Factible **(positivos diferentes de cero)**

$$Ax = \begin{bmatrix} 1 & 2 & -1 & 0 \\ 2 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 6 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \end{bmatrix} = b.$$

¿A cuál vértice ir a continuación? Se requiere moverse a lo largo de una arista hacia un vértice adyacente. Debido a que los dos vértices son vecinos, $m - 1$ variables básicas permanecen siendo básicas. *Sólo uno de los* 6s *se vuelve libre* (cero). Al mismo tiempo, *una variable se mueve desde cero para convertirse en básica*. Las otras $m - 1$ componentes básicas (en este caso, los otros 6) cambian pero siguen siendo positivas. La elección de arista (consulte el ejemplo 2 a continuación) decide qué variable sale de la base y cuál entra. Las variables básicas se calculan al resolver $Ax = b$. Las componentes libres de x se igualan a cero.

Ejemplo 2 Una variable que entra y una variable que sale producen un desplazamiento hacia un nuevo vértice.

$$\text{Minimizar} \quad 7x_3 - x_4 - 3x_5 \quad \text{sujeto a} \quad \begin{aligned} x_1 \quad + x_3 + 6x_4 + 2x_5 &= 8 \\ x_2 + x_3 \quad + 3x_5 &= 9. \end{aligned}$$

Se empieza desde el vértice en que $x_1 = 8$ y $x_2 = 9$ son las variables básicas. En ese vértice, $x_3 = x_4 = x_5 = 0$. Esto es factible, aunque el costo puede no ser el mínimo. Sería imprudente hacer positivo a x_3, ya que su coeficiente de costo es $+7$ y se está tratando de minimizar el costo. Se escoge x_5 porque tiene el coeficiente de costo más negativo: -3. ***La variable de entrada es x_5.***

Con x_5 que entra a la base, deben salir x_1 o x_2. En la primera ecuación, x_5 se incrementa y x_1 se disminuye mientras se mantiene $x_1 + 2x_5 = 8$. Luego, x_1 se lleva hasta cero cuando x_5 llega a 4. En la segunda ecuación se mantiene $x_2 + 3x_5 = 9$. Aquí x_5 sólo puede crecer hasta 3. Ir más allá haría negativo a x_2, de modo que ***la variable de salida es x_2.*** El nuevo vértice tiene $x = (2, 0, 0, 0, 3)$. *El costo ha bajado a* -9.

Forma rápida En $Ax = b$, los miembros derechos divididos entre los coeficientes de la variable de entrada son $\frac{8}{2}$ y $\frac{9}{3}$. La razón más pequeña, $\frac{9}{3}$ indica cuál es la variable que llega primero a cero, por lo que debe salir. Sólo se consideran razones positivas, porque si el coeficiente de x_5 fuese -3, entonces incrementar x_5 en realidad *haría crecer* a x_2. (En $x_5 = 10$, la segunda ecuación daría $x_2 = 39$.) ***La razón $\frac{9}{3}$ indica que la segunda variable sale.*** También proporciona $x_5 = 3$.

Si todos los coeficientes de x_5 hubieran sido negativos, se tendría un caso no acotado: es posible hacer arbitrariamente grande a x_5, y disminuir el costo hacia $-\infty$.

El paso actual termina en el nuevo vértice $x = (2, 0, 0, 0, 3)$. El paso siguiente es fácil sólo si las variables básicas x_1 y x_5 se mantienen por sí mismas (como originalmente fue el caso con x_1 y x_2). En consecuencia, se "pivotea" sustituyendo $x_5 = \frac{1}{3}(9 - x_2 - x_3)$ en la función de costo y en la primera ecuación. El nuevo problema, empezando desde el nuevo vértice, es:

$$\begin{aligned} &\text{Minimizar el costo} && 7x_3 - x_4 - (9 - x_2 - x_3) = x_2 + 8x_3 - x_4 - 9 \\ &\text{con restricciones} && x_1 - \tfrac{2}{3}x_2 + \tfrac{1}{3}x_3 + 6x_4 \qquad = 2 \\ & && \tfrac{1}{3}x_2 + \tfrac{1}{3}x_3 \qquad + x_5 = 3. \end{aligned}$$

El paso siguiente es fácil ahora. El único coeficiente negativo -1 en el costo hace que la variable de entrada sea x_4. Las razones de $\frac{2}{6}$ y $\frac{3}{0}$, los miembros derechos divididos entre la columna de x_4, hace que x_1 sea la variable de salida. El nuevo vértice es $x^* = (0, 0, 0, \frac{1}{3}, 3)$. El nuevo costo $-9\frac{1}{3}$ es el mínimo.

En un problema grande, una variable de salida puede volver a entrar a la base después. Pero el costo sigue bajando —excepto en un caso degenerado— de modo que las m variables básicas no pueden ser las mismas de antes. ¡Ningún vértice se visita dos veces! El método simplex debe terminar en el vértice óptimo (o en $-\infty$ si ocurre que el costo no está acotado). Lo extraordinario es la rapidez con que se encuentre x^*.

Resumen Los coeficientes de costo 7, -1, -3 en el primer vértice y 1, 8, -1 en el segundo vértice decidieron las variables de entrada. (Estos números van en r, el vector crucial que se define a continuación. El método se detiene cuando todos son positivos). Las razones decidieron las variables de salida.

Observación sobre la degeneración Un vértice es *degenerado* si más de las n componentes de costumbre de x son cero. Por el vértice pasan más de n planos, de modo que ocurre que una variable básica se hace cero. Las razones que determinan la variable de salida incluyen ceros, y la base podría cambiar sin realmente moverse del vértice. Teóricamente, es posible permanecer en un vértice y dar vueltas por siempre en la elección de la base.

Por fortuna, esta situación no ocurre. Es tan rara que los códigos comerciales la ignoran. Lamentablemente, los casos degenerados son bastante comunes en aplicaciones: si el costo se imprime después de cada paso simplex, se observa que se repite varias veces antes de que el método simplex encuentre un buen vértice. Luego, el costo decrece otra vez.

La tabla (tableau)

Cada paso simplex implica decisiones seguidas por operaciones en los renglones —es necesario escoger a las variables de entrada y de salida, y es necesario hacer que vayan y vuelvan. Una forma de organizar el paso es escribir A, b, c en una matriz grande, o ***tabla***:

$$\textbf{La tabla es de } m+1 \textbf{ por } m+n+1 \qquad T = \begin{bmatrix} A & b \\ c & 0 \end{bmatrix}.$$

Al inicio, las variables básicas pueden mezclarse con las variables libres. Volviendo a numerar en caso de ser necesario, *suponga que* $x_1, \ldots, x_m$ *son las variables básicas (distintas de cero) en el vértice actual*. Las m primeras columnas de A forman una matriz cuadrada B (la *matriz base* para ese vértice). Las n últimas columnas proporcionan una matriz N de m por n. El vector de costo se separa en $[c_B\ c_N]$, y la incógnita x lo hace en (x_B, x_N).

En el vértice, las variables libres son $x_N = 0$. Ahí, $Ax = b$ se convierte en $Bx_B = b$:

$$\textbf{Tabla en el vértice} \qquad T = \left[\begin{array}{c:c:c} B & N & b \\ \hdashline c_B & c_N & 0 \end{array}\right] \quad x_N = 0 \quad x_B = B^{-1}b \ \textbf{ costo } = c_B B^{-1} b.$$

Las variables básicas quedan solas cuando la eliminación multiplica por B^{-1}:

$$\textbf{Tabla reducida} \qquad T' = \left[\begin{array}{c:c:c} I & B^{-1}N & B^{-1}b \\ \hdashline c_B & c_N & 0 \end{array}\right].$$

Para alcanzar la ***forma escalonada totalmente reducida*** $R = \text{rref}(T)$, c_B multiplicado por el renglón en bloque superior se resta del renglón inferior:

$$\textbf{Totalmente reducida} \qquad R = \left[\begin{array}{c:c:c} I & B^{-1}N & B^{-1}b \\ \hdashline 0 & c_N - c_B B^{-1} N & -c_B B^{-1} b \end{array}\right].$$

A continuación se revisará el significado de cada elemento de esta tabla, y se llama la atención hacia el ejemplo 3 (a continuación, con números). He aquí el álgebra:

$$\textbf{Restricciones} \quad x_B + B^{-1}Nx_N = B^{-1}b \qquad \textbf{Vértice} \quad x_B = B^{-1}b, \quad x_N = 0. \tag{1}$$

El costo $c_Bx_B + c_Nx_N$ se ha convertido en

$$\textbf{Costo} \quad cx = (c_N - c_B B^{-1} N)x_N + c_B B^{-1} b \ \ \textbf{Costo en este vértice} = c_B B^{-1} b. \tag{2}$$

Toda cantidad importante aparece en la tabla completamente reducida R. Es posible decidir si el vértice es óptimo al observar a $r = c_N - c_B B^{-1} N$ en medio del renglón inferior. **Si cualquier elemento en r es negativo, sigue siendo posible reducir el costo.** Es posible hacer negativo a rx_N, al principio de la ecuación (2), incrementando una componente de x_N. Este es el paso siguiente. Pero si $r \geq 0$, entonces se ha encontrado el mejor vértice. Esta es la ***prueba de detención***, o ***condición de optimalidad***:

> **8B** El vértice es óptimo cuando $r = c_N - c_B B^{-1} N \geq 0$. Su costo es $c_B B^{-1} b$. Las componentes negativas de r corresponden a aristas en las cuales el costo disminuye. **La variable x_i corresponde a la componente más negativa de r.**

Las componentes negativas de r son los ***costos reducidos***: el costo en c_N por usar una variable *menos lo que ahorra*. El cálculo de r se denomina ***poner precio*** a las variables. Si el costo directo (en c_N) es menor que el ahorro (de reducir las variables básicas), entonces $r_i < 0$, y pagará por incrementar esa variable básica.

Suponga que el costo reducido más negativo es r_i. Entonces, la i-ésima componente de x_N es la ***variable de entrada***, que crece desde cero hasta un valor positivo α en el vértice siguiente (el final de la arista).

A medida que se incrementa x_i, otras componentes de x pueden decrecer (para mantener $Ax = b$). La primera x_k que llega a cero se convierte en la ***variable de salida***: cambia de básica a libre. *El siguiente vértice se alcanza cuando una componente de x_B cae a cero.*

El nuevo vértice es factible porque sigue teniéndose $x \geq 0$. Es básico porque de nuevo se tienen n componentes cero. La i-ésima componente de x_N pasó de cero a α. La k-ésima componente de x_B cayó a cero (las otras componentes de x_B siguen siendo positivas). La x_k de salida que cae a cero es la que proporciona la razón mínima en la ecuación (3):

> **8C** Suponga que x_i es la variable de entrada y que u es la columna i de N:
>
> $$\textbf{En el nuevo vértice} \quad x_i = \alpha = \text{razón más pequeña } \frac{(B^{-1}b)_j}{(B^{-1}u)_j} = \frac{(B^{-1}b)_k}{(B^{-1}u)_k}. \tag{3}$$
>
> Este mínimo sólo se toma sobre componentes positivas de $B^{-1}u$. La k-esima columna de B sale de la base (x_k se vuelve cero) y entra la nueva columna u.

$B^{-1}u$ es la columna de $B^{-1}N$ en la tabla reducida R, arriba del elemento más negativo en el renglón inferior r. Si $B^{-1}u \leq 0$, el siguiente vértice está infinitamente lejos y el costo mínimo es $-\infty$ (esto no ocurre aquí). Nuestro ejemplo va del vértice P a Q, y de nuevo empieza en Q.

Ejemplo 3 La función de costo original $x + y$ y las restricciones $Ax = b = (6, 6)$ proporcionan

$$\begin{bmatrix} A & b \\ c & 0 \end{bmatrix} = \left[\begin{array}{cccc:c} 1 & 2 & -1 & 0 & 6 \\ 2 & 1 & 0 & -1 & 6 \\ \hdashline 1 & 1 & 0 & 0 & 0 \end{array}\right].$$

En el vértice P en la figura 8.3, $x = 0$ corta a $2x + y = 6$. Para estar organizados, se intercambian las columnas 1 y 3 con la finalidad de escribir las variables básicas antes que las variables libres:

$$\textbf{Tabla en } P \qquad T = \left[\begin{array}{cc:cc:c} -1 & 2 & 1 & 0 & 6 \\ 0 & 1 & 2 & -1 & 6 \\ \hdashline 0 & 1 & 1 & 0 & 0 \end{array}\right].$$

Luego, la eliminación multiplica el primer renglón por -1, para obtener un pivote unitario, y usa el segundo renglón para producir ceros en la segunda columna:

$$\textbf{Totalmente reducida en } P \qquad R = \left[\begin{array}{cc:cc:c} 1 & 0 & 3 & -2 & 6 \\ 0 & 1 & 2 & -1 & 6 \\ \hdashline 0 & 0 & -1 & 1 & -6 \end{array}\right].$$

Primero observe a $r = [-1\ 1]$ en el renglón inferior. Tiene un elemento negativo en la columna 3, de modo que la tercera variable entra a la base. El vértice actual P y su costo $+6$ no son óptimos. La columna arriba de ese elemento negativo es $B^{-1}u = (3, 2)$; sus razones con la última columna son $\frac{6}{3}$ y $\frac{6}{2}$. Debido a que la primera razón es menor, la primera incógnita w (y la primera columna de la tabla) es obligada a abandonar la base. En la figura 8.3, se efectúa un movimiento a lo largo del conjunto factible del vértice P al vértice Q.

La nueva tabla intercambia las columnas 1 y 3, y al pivotear por eliminación se obtiene

$$\left[\begin{array}{cc:cc:c} 3 & 0 & 1 & -2 & 6 \\ 2 & 1 & 0 & -1 & 6 \\ \hdashline -1 & 0 & 0 & 1 & -6 \end{array}\right] \rightarrow \left[\begin{array}{cc:cc:c} 1 & 0 & \frac{1}{3} & -\frac{2}{3} & 2 \\ 0 & 1 & -\frac{2}{3} & \frac{1}{3} & 2 \\ \hdashline 0 & 0 & \frac{1}{3} & \frac{1}{3} & -4 \end{array}\right].$$

En esa nueva tabla en Q, $r = \left[\frac{1}{3} \quad \frac{1}{3}\right]$ es positivo. ***Se ha pasado la prueba de detención***. El vértice $x = y = 2$ y su costo $+4$ son óptimos.

Organización de un paso simplex

A continuación, la geometría del método simplex se representa en álgebra: Los "vértices" son "soluciones factibles básicas". El vector r y la razón α son decisivos. Su cálculo constituye el corazón del método simplex, y puede organizarse de tres formas distintas:

1. En una tabla, como arriba.
2. Actualizando B^{-1} cuando la columna u tomada de N sustituye a la columna k de B.
3. Calculando $B = LU$, y actualizando estos factores LU en vez de hacerlo con B^{-1}.

Esta lista constituye una breve historia del método simplex. En algunas formas, la etapa más fascinante es la primera —la *tabla*— que dominó el tema por muchos años. Para la mayoría de nosotros, trajo un aura de misterio a la programación lineal, principalmente porque se las arregló para evitar casi por completo la notación matricial (¡mediante el dispositivo habilidoso de escribir por completo todas las matrices!) Para efectos computacionales (excepto para pequeños problemas en libros de texto), la época de la tabla ya terminó.

Para ver por qué, recuerde que después de que el coeficiente más negativo en r indica cuál columna u entra a la base, no se utilizará ninguna de las otras columnas arriba de r. *Fue una pérdida de tiempo calcularlas*. En un problema más grande, cientos de columnas se calculan una y otra vez, justo para esperar su turno de entrar a la base. Esto permite que la teoría haga por completo las eliminaciones y alcanzar R. Sin embargo, en la práctica no es posible justificar esto.

Es más rápido, y al final más simple, ver cuáles cálculos son realmente necesarios. Cada paso simplex intercambia una columna de N por una columna de B. Estas columnas son decididas por r y α. Este paso comienza con la matriz base actual B y la solución actual $x_B = B^{-1}b$.

Un paso del método simplex

1. Se calculan el vector renglón $\lambda = c_B B^{-1}$ y los costos reducidos $r = c_N - \lambda N$.
2. Si $r \geq 0$, detener el procedimiento: la solución actual es óptima. En caso contrario, si r_i es la componente más negativa, para entrar a la base se escoge u = columna i de N.
3. Se calculan las razones de $B^{-1}b$ a $B^{-1}u$, admitiendo sólo componentes positivas de $B^{-1}u$. (Si $B^{-1}u < 0$, entonces el costo mínimo es $-\infty$). Cuando la razón más pequeña ocurre en la componente k, de la base sale la k-ésima columna de la B actual.
4. Se actualiza B, B^{-1}, o LU, así como la solución $x_B = B^{-1}b$. Se vuelve al paso **1**.

Algunas veces, lo anterior se denomina ***método simplex revisado*** para distinguirlo de las operaciones en la tabla. En realidad es el método simplex en sí, reducido.

Este análisis termina una vez que se ha decidido cómo calcular los pasos **1**, **3**, y **4**:

$$\lambda = c_B B^{-1}, \quad v = B^{-1}u, \quad \text{y} \quad x_B = B^{-1}b. \tag{4}$$

La forma más conocida es trabajar directamente con B^{-1}, calculándola explícitamente en el primer vértice. En vértices subsecuentes, el paso de pivoteo es sencillo. Cuando la columna k de la matriz identidad se sustituye por u, la columna k de B^{-1} se sustituye por $v = B^{-1}u$. Para reconvertir la matriz identidad, la eliminación multiplica la B^{-1} anterior por

$$E^{-1} = \begin{bmatrix} 1 & & v_1 & & \\ & \ddots & \vdots & & \\ & & v_k & & \\ & & \vdots & \ddots & \\ & & v_n & & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & & -v_1/v_k & & \\ & \ddots & \vdots & & \\ & & 1/v_k & & \\ & & \vdots & \ddots & \\ & & -v_n/v_k & & 1 \end{bmatrix} \tag{5}$$

En muchos códigos simplex se utiliza la ***forma de producto de la inversa***, que ahorra estas simples matrices E^{-1} en vez de actualizar directamente a B^{-1}. Cuando es necesario, se aplican a b y c_B. A intervalos regulares (quizá cada 40 pasos simplex), se vuelve a calcular B^{-1} y la E^{-1} se borra. En el problema 9 se comprueba la ecuación (5), al final de esta sección.

En un método más reciente se usan los métodos normales del álgebra lineal numérica, considerando a la ecuación (4) como tres ecuaciones que comparten la misma matriz B:

$$\lambda B = c_B, \quad Bv = u, \quad Bx_B = b. \tag{6}$$

La factorización de costumbre $B = LU$ (o $PB = LU$, con intercambios de renglones por razones de estabilidad) conduce a las tres soluciones. L y U pueden actualizarse, en vez de calcularlas de nuevo.

Queda una pregunta: *¿Cuántos pasos simplex es necesario efectuar?* Es imposible contestar de antemano. La experiencia indica que el método toca sólo alrededor de $3m/2$ vértices distintos, lo cual significa un conteo de operaciones aproximado de m^2n. Esto es comparable con la eliminación normal para $Ax = b$, y es la razón del éxito del método simplex. Sin embargo, las matemáticas muestran que la longitud de la ruta no siempre puede acotarse por cualquier múltiplo fijo o potencia de m. Los peores conjuntos factibles (Klee y Minty inventaron un cubo lobulado) pueden obligar a que el método simplex intente cada vértice, a costo exponencial.

Fue el ***método de Khachian*** lo que demostró que la programación lineal podía resolverse en tiempo polinomial.* Su algoritmo permaneció *dentro* del conjunto factible, y capturó a x^* en una serie de elipsoides que se contraen. La programación lineal está en la clase agradable ***P***, no en la temible clase ***NP*** (como el problema del agente de ventas viajero). Para problemas ***NP***, se considera (aunque no está demostrado) que todos los algoritmos determinísticos deben terminar en un tiempo interminablemente exponencial, en el peor de los casos.

Todo este tiempo, el método simplex ha hecho su trabajo —en un *tiempo* medio del que ya se ha demostrado (para variantes del método de costumbre) que es polinomial. Por

* El número de operaciones está acotado por potencias de m y n, como en la eliminación. Para programación entera y factorización en primos, todos los algoritmos conocidos pueden ser exponencialmente largos. La célebre conjetura "$P \neq NP$" establece que tales problemas no pueden tener algoritmos polinomiales.

alguna razón, oculta en la geometría de poliedros de muchas dimensiones, los malos conjuntos factibles son raros y el método simplex tiene suerte.

Método de Karmarkar

Ahora llegamos al evento más sensacional que ha ocurrido en la historia reciente de la programación lineal. Karmarkar propuso un método basado en dos conceptos simples, y en sus experimentos derrotó al método simplex. Tanto la elección del problema como los detalles del código son cruciales, y el debate sigue vigente. Sin embargo, las ideas de Karmarkar eran tan naturales y se ajustaban tan perfectamente al marco de referencia del álgebra lineal aplicada, que pueden explicarse en unos cuantos párrafos.

La primera idea es comenzar desde un punto que ***esté dentro del conjunto factible***: se supondrá que es $x^0 = (1, 1, \ldots, 1)$. Debido a que el costo es cx, la mejor *dirección para reducir el costo* es hacia $-c$. Normalmente, lo anterior nos saca del conjunto factible; moverse en esa dirección no preserva $Ax = b$. Si $Ax^0 = b$ y $Ax^1 = b$, entonces $\Delta x = x^1 - x^0$ debe satisfacer $A\Delta x = 0$. ***El paso Δx debe estar en el espacio nulo de A.*** En consecuencia, $-c$ *se proyecta* sobre el espacio nulo, con la finalidad de encontrar la dirección factible más próxima a la mejor dirección. Este es un paso natural pero costoso en el método de Karmarkar.

El paso Δx es un múltiplo de la proyección $-Pc$. Mientras más grande sea el paso, más se reduce el costo, aunque no es posible salir del conjunto factible. El múltiplo de $-Pc$ se escoge de modo que x^1 esté próximo, aunque un *poco adentro*, de la frontera en la cual una componente de x llega a cero.

Así se completa la primera idea: la proyección que proporciona el *descenso factible más pronunciado*. El segundo paso requiere una nueva idea, ya que continuar en la misma dirección es inútil.

La sugerencia de Karmarkar es ***transformar x^1 de vuelta a (1, 1, . . . , 1) en el centro.*** Su cambio de variables no era lineal, pero la transformación más simple es justo un ***reescalamiento*** por una matriz diagonal D. Luego ya hay espacio para moverse. El reescalamiento desde x hasta $X = D^{-1}x$ cambia las restricciones y el costo:

$$Ax = b \text{ se vuelve } ADX = b \qquad c^{\mathrm{T}}x \text{ se vuelve } c^{\mathrm{T}}DX.$$

En consecuencia, *la matriz AD toma el sitio de A, y el vector $c^{\mathrm{T}}D$ toma el lugar de c^{T}*. El segundo paso proyecta la nueva c sobre el espacio nulo de la nueva A. Todo el trabajo está en esta proyección, para resolver las ecuaciones normales ponderadas:

$$(AD^2A^{\mathrm{T}})y = AD^2c. \tag{7}$$

La forma normal para calcular y es por eliminación. El proceso de Gram-Schmidt ortogonaliza las columnas de DA^{T}, lo cual puede resultar costoso (aunque facilita el resto de los cálculos). El método favorito para problemas ralos grandes es el ***método del gradiente conjugado***, que proporciona la respuesta exacta más lentamente que la eliminación, aunque puede aplicarse parcialmente y luego detenerse. A mitad de la eliminación no es posible detenerse.

Así como otras ideas nuevas en la computación científica, el método de Karmarkar tuvo éxito en algunos problemas y en otros no. La idea subyacente fue analizada y mejorada. Un éxito primordial lo constituyen **métodos de punto interior** (que permanecen dentro del conjunto factible) más recientes, que se mencionan en la siguiente sección. Y el método simplex sigue siendo tremendamente valioso, como todo el tema de la programación lineal, que fue descubierta siglos después que $Ax = b$, aunque comparte las ideas fundamentales del álgebra lineal. De estas ideas, la que va más lejos es la de dualidad, que se presenta a continuación.

Conjunto de problemas 8.2

1. Minimizar $x_1 + x_2 - x_3$, sujeto a

$$\begin{aligned} 2x_1 - 4x_2 + x_3 + x_4 \quad &= 4 \\ 3x_1 + 5x_2 + x_3 \quad + x_5 &= 2. \end{aligned}$$

¿Cuál de x_1, x_2, x_3 debe entrar a la base, y cuál de x_4, x_5, debe salir? Calcule el nuevo par de variables básicas, y encuentre el costo en el nuevo vértice.

2. Después del paso simplex precedente, prepare y decida el paso siguiente.

3. En el ejemplo 3, suponga que el costo es $3x + y$. Con reagrupamiento, el vector de costo es $c = (0, 1, 3, 0)$. Demuestre que $r \geq 0$ y, en consecuencia, que el vértice P es óptimo.

4. Suponga que la función de costo en el ejemplo 3 es $x - y$, de modo que después de un reagrupamiento $c = (0, -1, 1, 0)$ en el vértice P. Calcule r y decida qué columna u debe ser elemento de la base. Entonces calcule $B^{-1}u$ y muestre desde su signo que nunca llegará a otro vértice. En la figura 8.3 estamos escalando el eje y y $x - y$ llega a $-\infty$.

5. De nuevo en el ejemplo 3, cambie el costo a $x + 3y$. Compruebe que el método simplex va de P a Q a R, y que el vértice R es óptimo.

6. *La Fase I encuentra una solución factible básica de $Ax = b$* (un vértice). Después de cambiar signos para hacer $b \geq 0$, considere el problema auxiliar de minimizar $w_1 + w_2 + \ldots + w_m$, sujeto a $x \geq 0$, $w \geq 0$, $Ax + w = b$. Siempre que $Ax = b$ tiene una solución nonegativa, el costo mínimo en este problema es cero, con $w^* = 0$.

a) Demuestre que, para este nuevo problema, el vértice $x = 0$, $w = b$ es tanto básico como factible. En consecuencia, *su* Fase I ya está establecida, y el método simplex puede proceder para encontrar el par óptimo x^*, w^*. Si $w^* = 0$, entonces x^* es el vértice requerido en el problema original.

b) Con $A = [1\ -1]$ y $b = [3]$, escriba el problema auxiliar, el vector de su Fase I $x = 0$, $w = b$, y su vector óptimo. Encuentre el vértice del conjunto factible $x_1 - x_2 = 3$, $x_1 \geq x_2 \geq 0$, y trace una figura de este conjunto.

7. Si quisiera maximizarse el costo en vez de minimizarlo (con $Ax = b$ y $x \geq 0$), ¿cuál sería la prueba de detención sobre r, y qué reglas escogerían la columna de N con la finalidad de hacerla básica y la columna de B para hacerla libre?

8. Minimice $2x_1 + x_2$, sujeto a $x_1 + x_2 \geq 4$, $x_1 + 3x_2 \geq 12$, $x_1 - x_2 \geq 0$, $x \geq 0$.

9. Compruebe la inversa en la ecuación (5), y demuestre que BE tiene $Bv = u$ en su k-ésima columna. Así, BE es la matriz básica correcta para la siguiente detención, $E^{-1}B^{-1}$ es su inversa, y E^{-1} actualiza correctamente la matriz básica.

10. Suponga que se quiere minimizar $cx = x_1 - x_2$, sujeto a

$$\begin{aligned} 2x_1 - 4x_2 + x_3 \quad &= 6 \\ 3x_1 + 6x_2 \quad + x_4 &= 12 \end{aligned} \qquad (\text{toda } x_1, x_2, x_3, x_4 \geq 0).$$

Empezando desde $x = (0, 0, 6, 12)$, ¿deben incrementarse x_1 o x_2 a partir de su valor actual de cero? ¿Cuánto puede incrementarse hasta que la ecuación obliga a cero a x_3 o x_4? En ese punto, ¿cuál es el nuevo x?

11. Para la matriz $P = I - A^T(AA^T)^{-1}A$, demuestre que si x está en el espacio nulo de A, entonces $Px = x$. El espacio nulo permanece sin cambio bajo esta proyección.

12. *a*) Minimice el costo $c^Tx = 5x_1 + 4x_2 + 8x_3$ sobre el plano $x_1 + x_2 + x_3 = 3$, probando los vértices P, Q, R, donde el triángulo está cortado por el requerimiento $x \geq 0$.

b) Proyecte $c = (5, 4, 8)$ sobre el espacio nulo de $A = [1\ 1\ 1]$, y encuentre el paso máximo s que mantiene nonegativo a $e - sPc$.

8.3 EL PROBLEMA DUAL

La eliminación puede resolver $Ax = b$, pero los cuatro subespacios fundamentales han demostrado que una comprensión diferente y más profunda es posible. Es exactamente lo mismo para la programación lineal. La mecánica del método simplex resuelve un programa lineal, aunque la dualidad se encuentra en el centro de la teoría subyacente. La introducción del problema dual es una idea elegante, que al mismo tiempo es fundamental para las aplicaciones. Se explicará tanto como entendemos.

La teoría empieza con el *problema original* dado:

Original (P) *Minimizar cx, sujeto a* $x \geq 0$ *y* $Ax \geq b$.

El problema dual comienza a partir de las mismas A, b, y c, e invierte todo. En el problema original, c está en la función de costo y b está en la restricción. En el problema dual, se cambian b y c. La incógnita dual y es un *vector renglón* con m componentes, y el conjunto factible tiene $yA \leq c$, en vez de $Ax \geq b$.

En corto, el dual de un problema mínimo es un problema máximo. Ahora $y \geq 0$:

Dual (D) *Maximizar yb, sujeto a* $y \geq 0$ *y* $yA \leq c$.

El dual de *este* problema es el problema mínimo original. Hay simetría completa entre los problemas dual y original. El método simplex es igualmente válido para una maximización; en todo caso, ambos problemas se resuelven a la vez.

Es necesario proporcionar una interpretación de todas estas inversiones. Aconsejan una competencia entre el minimizador y el maximizador. En el problema de la dieta, el minimizador cuenta con n alimentos (mantequilla de cacahuate y bistec, en la sección 8.1). Entran en el problema de la dieta en las cantidades (nonegativas) $x_1, \ldots, x_n$. Las restricciones representan *m vitaminas requeridas*, en vez de una restricción anterior de suficientes proteínas. El elemento a_{ij} mide la i-ésima vitamina en el j-ésimo alimento, y el i-ésimo renglón de $Ax \geq b$ obliga a que en la dieta se incluya por lo menos b_i de esa vitamina. Si c_j es el costo del j-ésimo alimento, entonces $c_1x_1 + \cdots + c_nx_n = cx$ es el costo de la dieta. Este es el costo que debe minimizarse.

En el dual, el farmacéutico vende píldoras de vitamina, a precios $y_i \geq 0$. Debido a que el alimento j contiene vitaminas en las cantidades a_{ij}, el precio del farmacéutico para el equivalente de vitaminas no puede exceder el precio c_j del tendero. Esta es la j-ésima restricción en $yA \leq c$. Al trabajar dentro de esta restricción sobre los precios de las vitaminas, el farmacéutico puede vender la cantidad requerida b_i de cada vitamina por un ingreso total de $y_1b_1 + \cdots + y_mb_m = yb$ a maximizar.

Los conjuntos factibles para los problemas original y dual se ven completamente diferentes. El primero es un subconjunto de $\mathbf{R}^n$, marcado por $x \geq 0$ y $Ax \geq b$. El segundo es un subconjunto de $\mathbf{R}^m$, determinado por $y \geq 0$ y A^T y c. Toda la teoría de la programación lineal radica en la relación que hay entre el original y el dual. A continuación se presenta el resultado fundamental:

8D Teorema de dualidad Cuando ambos problemas tienen vectores factibles, tienen x^* y y^* óptimos. **El costo mínimo cx^* es igual al ingreso máximo y^*b.**

Si los vectores óptimos no existen, hay dos posibilidades: ya sea que ambos conjuntos factibles son vacíos o uno es vacío y el otro problema no está acotado (el máximo es $+\infty$ o el mínimo es $-\infty$).

El teorema de dualidad establece la competencia entre el tendero y el farmacéutico. *El resultado siempre es un empate*. En la teoría de juegos se encuentra un "teorema minimax" semejante. El cliente no tiene ninguna razón económica para preferir las vitaminas a los alimentos, aun cuando el farmacéutico garantice igualar al tendero en cada alimento —e incluso venda más barato los alimentos costosos (como la mantequilla de cacahuate). Se demostrará que los alimentos costosos se mantienen fuera de la dieta original, de modo que la salida puede ser (y es) un empate.

Lo anterior puede parecer un punto muerto, aunque espero que el lector no se deje engañar. Los vectores óptimos contienen la información crucial. En el problema original, x^* indica al comprador qué comprar. En el dual, y^* fija los precios naturales (***precios imaginarios***) a los cuales la economía debería marchar. En tanto nuestro modelo lineal refleje la verdadera economía, x^* y y^* representan las decisiones esenciales que han de tomarse.

Se quiere demostrar que $cx^* = y^*b$. Puede parecer evidente que el farmacéutico pueda subir los precios y^* de las vitaminas para encontrar al comprador, aunque sólo una cuestión es verdaderamente clara: Debido a que cada alimento puede sustituirse por su equivalente vitamínico, sin incremento en el costo, todas las dietas alimentarias idóneas deben costar por lo menos tanto como las vitaminas. Esta desigualdad es unilateral: *precio del farmacéutico* $\leq$ *precio del tendero*. Se denomina ***dualidad débil***, y es fácil demostrarlo para cualquier programa lineal y su dual:

8E Si x y y son factibles en los problemas original y dual, entonces $yb \leq cx$.

Demostración Debido a que los vectores son factibles, satisfacen $Ax \geq b$ y $yA \leq c$. Debido a que la factibilidad también incluye $x \geq 0$ y $y \geq 0$, es posible tomar productos internos sin arruinar estas desigualdades (multiplicar por números negativos debe invertirlos):

$$yAx \geq yb \quad \text{y} \quad yAx \leq cx. \tag{1}$$

Debido a que los miembros izquierdos son idénticos, se tiene la dualidad débil $yb \leq cx$. ■

Esta desigualdad unilateral prohíbe la posibilidad de que ambos problemas sean no acotados. Si yb es arbitrariamente grande, una x factible contradice $yb \leq cx$. De manera semejante, si cx puede ir hacia $-\infty$, entonces el dual no puede admitir una y factible.

Igualmente importante, cualesquiera vectores que alcancen $yb = cx$ deben ser óptimos. En este punto el precio del tendero es igual al del farmacéutico. Una dieta alimentaria óptima y los precios óptimos de las vitaminas se reconocen por el hecho de que el

consumidor no tiene nada que elegir:

8F Si los vectores x y y son factibles y $cx = yb$, entonces x y y son óptimos.

Debido a que ningún y puede hacer yb más grande que cx, la y que alcanza este valor es óptima. De manera semejante, cualquier x que alcance el costo $cx = yb$ debe ser una x^* óptima.

A continuación se proporciona un ejemplo con dos alimentos y dos vitaminas. Observe que A^T aparece cuando se escribe el dual, ya que $yA \leq c$ para vectores renglón significa $A^T y^T \leq c^T$ para columnas.

Original Minimizar $x_1 + 4x_2$ sujeto a $x_1 \geq 0,\ x_2 \geq 0$

$$2x_1 + x_2 \geq 6$$
$$5x_1 + 3x_2 \geq 7.$$

Dual Maximizar $6y_1 + 7y_2$ sujeto a $y_1 \geq 0,\ y_2 \geq 0$

$$2y_1 + 5y_2 \leq 1$$
$$y_1 + 3x_2 \leq 4.$$

Solución $x_1 = 3$ y $x_2 = 0$ son factibles, con costo $x_1 + 4x_2 = 3$. En el dual, $y_1 = \frac{1}{2}$ y $y_2 = 0$ proporcionan el mismo valor $6y_1 + 7y_2 = 3$. Estos vectores deben ser óptimos.

Por favor, observe la situación detenidamente para ver lo que ocurre realmente en el momento en que $yb = cx$. Algunas de las restricciones de desigualdad son **estrictas**, lo cual significa que se cumple la igualdad. Otras restricciones son flojas, y la regla clave tiene sentido económico:

i) La dieta tiene $x_j^* = 0$ cuando el precio de los alimentos j se fija *por arriba* de su equivalente vitamínico.
ii) El precio es $y_i^* = 0$ cuando se tiene una *sobreoferta* de vitamina i en la dieta x^*.

En el ejemplo, $x_2 = 0$ porque el segundo alimento es demasiado costoso. Su precio excede el precio del farmacéutico, ya que $y_1 + 3y_2 \leq 4$ es una desigualdad estricta $\frac{1}{2} + 0 < 4$. De manera semejante, la dieta requería siete unidades de la segunda vitamina, aunque en realidad proporcionó $5x_1 + 3x_2 = 15$. Así, se encontró $y_2 = 0$, y esta vitamina es un *producto libre*. Puede verse cómo se ha completado la dualidad.

Estas ***condiciones de optimalidad*** pueden entenderse fácilmente en términos matriciales. A partir de la ecuación (1) se quiere $y^*Ax^* = y^*b$ en el óptimo. La factibilidad requiere $Ax^* \geq b$, y se buscan cualesquiera componentes *en las que fracase la igualdad*. Esto corresponde a una vitamina que esté sobreofrecida, de modo que su precio es $y_i^* = 0$.

Al mismo tiempo, se tiene $y^*A \leq c$. Todas las restricciones estrictas (alimentos costosos) corresponden a $x_j^* = 0$ (omisión en la dieta). Esta es la clave para $y^*Ax^* = cx^*$, que se requiere. Estas son las ***condiciones flojas complementarias*** de la programación lineal, y las ***condiciones de Kuhn-Tucker*** de programación no lineal:

8G Los vectores óptimos x^* y y^* satisfacen la **flojedad complementaria:**

$$\text{si } (Ax^*)_i > b_i \text{ entonces } y_i^* = 0 \qquad \text{si} (y^*A)_j < c_j \text{ entonces } x_j^* = 0. \tag{2}$$

A continuación se repite la demostración. Cualesquiera vectores factibles x y y satisfacen la dualidad débil:

$$yb \leq y(Ax) = (yA)x \leq cx. \tag{3}$$

Se requiere la igualdad, y sólo hay una forma en que y^*b puede ser igual a $y^*(Ax^*)$. *Cada vez que $b_i < (Ax^*)_i$, el factor y_i^* que multiplica a estas componentes debe ser cero.*

De manera semejante, la factibilidad proporciona $yAx \leq cx$. La igualdad se obtiene sólo cuando se cumple la segunda condición de flojera. Si hay un sobreprecio $(y^*A)_j < c_j$, debe cancelarse a través de la multiplicación por $x_j^* = 0$. Así, se queda con $y^*b = cx^*$ en la ecuación (3). Esta igualdad garantiza la optimalidad de x^* y y^*.

Demostración de la dualidad

La desigualdad unilateral $yb \leq cx$ fue fácil de demostrar; proporcionó una prueba rápida para vectores óptimos (la convierten en igualdad); y ahora ha proporcionado las condiciones de flojera en la ecuación (2). Lo único que no ha hecho es mostrar que $y^*b = cx^*$ realmente es posible. El teorema de dualidad no está completo sino hasta que se obtienen estos vectores óptimos.

Para obtener y^*, se regresa al método simplex, que ya ha calculado x^*. El problema consiste en mostrar que el método se detuvo en el momento indicado para el problema dual (aun cuando se construyó para resolver el problema original). Recuerde que las m desigualdades $Ax \geq b$ se cambiaron a ecuaciones, mediante la introducción de las variables flojas $w = Ax - b$:

$$\textbf{Factibilidad original} \qquad \begin{bmatrix} A & -I \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} = b \quad \text{y} \quad \begin{bmatrix} x \\ w \end{bmatrix} \geq 0. \tag{4}$$

Cada paso simplex escogió m columnas de la matriz larga $[A \ \ -I]$ para hacerlas básicas, y las desplazó (teóricamente) hacia el frente. Así se obtuvo $[B \ \ N]$. El mismo desplazamiento reordenó el largo vector de costo $[c \ \ 0]$ en $[c_B \ \ c_N]$. La condición de detención, que llevó a término al método simplex, era $r = c_N - c_B B^{-1} N \geq 0$.

Esta condición $r \geq 0$ se alcanzó finalmente, ya que el número de vértices es finito. En ese momento, el costo era lo más bajo posible:

$$\textbf{Costo mínimo} \qquad cx^* = \begin{bmatrix} c_B & c_N \end{bmatrix} \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix} = c_B B^{-1} b. \tag{5}$$

Si en el dual puede escogerse $y^* = c_B B^{-1}$, ciertamente se tiene $y^*b = cx^*$. El mínimo y el máximo son iguales. Debe demostrarse que esta y^* satisface las restricciones duales $yA \leq c$ y $y \geq 0$:

$$\textbf{Factibilidad dual} \qquad y \begin{bmatrix} A & -I \end{bmatrix} \leq \begin{bmatrix} c & 0 \end{bmatrix}. \tag{6}$$

Cuando el método simplex vuelve a mezclar la gran matriz y el vector para poner primero las variables básicas, esto reagrupa las restricciones en la ecuación (6) como

$$y \begin{bmatrix} B & N \end{bmatrix} \leq \begin{bmatrix} c_B & c_N \end{bmatrix}. \tag{7}$$

Para $y^* = c_B B^{-1}$, la primera mitad es una igualdad y la segunda es $c_B B^{-1} N \leq c_N$. ¡Esta es la condición de detención $r \geq 0$ que se sabe debe cumplirse! En consecuencia, la y^* es factible, y *se ha demostrado el teorema de dualidad.* Al localizar la matriz crítica B de m por m, que es no singular en tanto esté prohibida la degeneración, el método simplex ha producido la y^* óptima, así como x^*.

Precios imaginarios

En cálculo, todo mundo conoce la condición para un máximo o un mínimo: *las primeras derivadas son cero*. Sin embargo, este hecho es transformado completamente por las restricciones. El ejemplo más sencillo es la recta $y = x$. Su derivada nunca es cero, el cálculo parece inútil, y ciertamente la mayor y ocurre al final del intervalo. ¡Esta es exactamente la situación en programación lineal! Hay más variables, y un intervalo se sustituye por un conjunto factible, aunque el máximo sigue encontrándose en un vértice del conjunto factible (con sólo m componentes diferentes de cero).

El problema en programación lineal es localizar ese vértice. Para este efecto, el cálculo no es completamente inútil. Lejos de ello, porque los "multiplicadores de Lagrange" harán cero las derivadas en el máximo y en el mínimo. *Las variables duales y son exactamente los multiplicadores de Lagrange*. Y responden la pregunta clave: **¿Cómo cambia el costo mínimo $cx^* = y^*b$, si se cambian b o c?**

Esta es una pregunta en ***análisis de sensibilidad***. Permite obtener información adicional del problema dual. Para un economista o un ejecutivo, estas preguntas sobre ***costo marginal*** son las más importantes.

Si se permiten grandes cambios en b o en c, la situación se comporta de manera bastante irregular. Cuando aumenta el precio de los huevos, hay un punto en el que éstos desaparecen de la dieta. La variable x_{huevo} salta de básica a libre. Para seguirla correctamente, es necesario introducir programación "paramétrica". Pero si los cambios son pequeños, ***el vértice que era óptimo permanece siendo óptimo***. La elección de las variables básicas no cambia; B y N siguen siendo las mismas. Geométricamente, el conjunto factible se ha desplazado un poco (al cambiar b), y los planos que lo cortan se han inclinado (al cambiar c). Cuando estos cambios son pequeños, el contacto ocurre en el mismo vértice (que se ha movido ligeramente).

Al final del método simplex, cuando se conocen las variables básicas idóneas, las m columnas correspondientes de A constituyen la matriz base B. En ese vértice, un desplazamiento de tamaño Δb modifica el costo mínimo por $y^*\Delta b$. ***La solución dual y^* proporciona la razón de cambio del costo mínimo*** (su derivada) ***respecto a cambios en b***. Las componentes de y^* son los ***precios imaginarios***. Si el requerimiento para una vitamina sube por Δ y el precio del farmacéutico es y_1^*, entonces el costo de la dieta (del farmacéutico o del tendero) sube por $y_1^*\Delta$. En el caso en que y_1^* es cero, esa vitamina es un *producto libre* y el pequeño cambio no tiene ningún efecto. La dieta ya contenía más que b_1.

Ahora se plantea una pregunta diferente. Suponga que se insiste en que la dieta contenga una pequeña cantidad digerible de huevo. La condición $x_{\text{huevo}} \geq 0$ cambia a $x_{\text{huevo}} \geq \delta$. ¿Cómo modifica esto al costo?

Si en la dieta x^* hubiera huevo, no habría ningún cambio. Pero si $x^*_{\text{huevo}} = 0$, habrá un costo extra por agregar la cantidad δ. El incremento no será en todo el precio $c_{\text{huevo}}\delta$, ya que es posible reducirlo de los otros alimentos. El ***costo reducido*** de los huevos es su propio precio, *menos* el precio que se paga por el equivalente en alimentos más baratos. Para calcularlo, se regresa a la ecuación (2) de la sección 8.2:

$$\text{costo} = (c_N - c_B B^{-1}N)x_N + c_B B^{-1}b = rx_N + c_B B^{-1}b.$$

Si el huevo es la primera variable libre, entonces el incremento de la primera componente de x_N a δ aumenta el costo por $r_1\delta$. *El verdadero costo del huevo es r_1*. Este es el cambio en el costo de la dieta cuando la cota inferior cero (restricción de no negatividad) se mueve hacia arriba. Se sabe que $r \geq 0$, y la economía indica lo mismo: el costo reducido de los huevos no puede ser negativo, o habrían entrado a la dieta.

Métodos de puntos interiores

El método simplex se mueve a lo largo de los vértices del conjunto factible, para llegar finalmente al vértice óptimo x^*. Los métodos de puntos interiores empiezan ***dentro*** del conjunto factible (donde todas las restricciones son *des*igualdades). Estos métodos esperan moverse directamente a x^* (y también encontrar y^*). Una vez que están muy próximos a la respuesta, se detienen.

Una forma de permanecer dentro es colocando una barrera en la frontera. Agregar un costo extra en la forma de un logaritmo que se infla cuando cualquier variable x o cualquier variable floja $w = Ax - b$ toca cero. El número θ es un pequeño parámetro a escoger:

$$\textbf{Problema con barrera } P(\theta) \qquad \text{Minimizar } cx - \theta\left(\sum_1^n \ln x_i + \sum_1^m \ln w_i\right). \tag{8}$$

Este costo es no lineal (aunque la programación lineal de hecho ya no es lineal, a partir de las desigualdades). La notación es más simple si el vector largo (x, w) vuelve a identificarse como x y $[A \ -I]$ se vuelve a identificar como A. Ahora, las restricciones originales son $x \geq 0$ y $Ax = b$. La suma de $\ln x_i$ en la barrera pasa ahora a $m + n$.

Las restricciones duales son $yA \leq c$. (Cuando se tiene $Ax = b$ en el original no se requiere $y \geq 0$). La variable floja es $s = c - yA$, con $s \geq 0$. ¿Cuáles son las condiciones de Kuhn-Tucker para que x y y sean las x^* y y^* óptimas? Junto con las restricciones se requiere dualidad: $cx^* = y^*b$.

Incluyendo la barrera se obtiene un *problema aproximado* $P(\theta)$. Para sus condiciones de Kuhn-Tucker de optimalidad, la derivada de $\ln x_i$ proporciona $1/x_i$. Si a partir de estos números positivos x_i se crea una matriz diagonal X, y se usa $e = [1 \ \ldots \ 1]$ para el renglón vector de $n + m$ unos, entonces la optimalidad en $P(\theta)$ es como sigue:

Original (vectores columna)	$Ax = b$ con $x \geq 0$	(9*a*)
Dual (vectores renglón)	$yA + \theta eX^{-1} = c$	(9*b*)

Cuando $\theta \to 0$, es de esperar que las x y y óptimas tiendan a x^* y y^* para el problema original sin barrera, y que θeX^{-1} permanezca no negativo. El plan es resolver las ecuaciones (9a y 9*b*) con barreras cada vez más pequeñas, dadas por el tamaño de θ.

En realidad, estas ecuaciones no lineales se resuelven aproximadamente con el método de Newton (lo cual significa que son linealizadas). El término no lineal es $s = \theta eX^{-1}$. Para evitar $1/x_i$, lo anterior se vuelve a escribir como $sX = \theta e$. Al crear la matriz diagonal S a partir de s, ésta es $eSX = \theta e$. Si e, y, c, y s se cambian a vectores columna, y se traspone, ahora la optimalidad consta de tres partes:

Original	$Ax = b, \ x \geq 0.$	(10*a*)
Dual	$A^{\mathrm{T}}y + s = c.$	(10*b*)
No lineal	$XSe - \theta e = 0.$	(10*c*)

El método de Newton requiere un paso Δx, Δy, Δs, a partir de las x, y, s, actuales. (Éstas resuelven las ecuaciones (10*a*) y (10*b*), pero no la (10*c*).) Al ignorar el término de segundo orden $\Delta X \Delta S e$, ¡las correcciones provienen de ecuaciones lineales!

	$A\,\Delta x = 0.$	(11*a*)
Paso de Newton	$A^{\mathrm{T}}\Delta y + \Delta s = 0.$	(11*b*)
	$S\,\Delta x + X\,\Delta s = \theta e - XSe.$	(11*c*)

Las notas de Robert Freund para su curso en el MIT afianzan la razón de convergencia (cuadrática) y la complejidad computacional de este algoritmo. Sin importar las dimensiones m y n, la brecha de dualidad sx suele estar abajo de 10^{-8} luego de entre 20 a 80 pasos de Newton. Este algoritmo se utiliza casi "como es" en software comercial de punto interior, así como para una amplia variedad de problemas de optimización no lineal.

Teoría de las desigualdades

La dualidad puede estudiarse en más de una forma. Rápidamente se demostró $yb \leq cx$, y luego se aplicó el método simplex para obtener la igualdad. Esta es una *prueba constructiva*; x^* y y^* se calculan en realidad. A continuación se aborda brevemente un método diferente, en el que se omite el algoritmo simplex, y se analiza más directamente la geometría. Considero que los conceptos clave son igualmente claros (de hecho, quizá más claros) si se omiten algunos detalles.

La mejor ilustración de este método se encuentra en el Teorema Fundamental del Álgebra Lineal. El problema en el capítulo 2 consistía en encontrar b en el espacio columna de A. Después de la eliminación y los cuatro subespacios, la cuestión de resolubilidad fue respondida en una forma completamente distinta por el problema 11 en la sección 3.1:

8H $Ax = b$ tiene una solución *o* existe una y tal que $yA = 0$ y $yb \neq 0$.

Este es el ***teorema de la alternativa***, porque encontrar tanto a x como a y es imposible: si $Ax = b$ entonces $yAx = yb \neq 0$, lo cual contradice a $yAx = 0x = 0$. En el lenguaje de los subespacios, b está en el espacio columna, o tiene una componente adherida en el espacio nulo izquierdo. Esta componente es la y que se busca.

Para desigualdades, se desea encontrar un teorema que sea exactamente del mismo tipo. Se empieza con el mismo sistema $Ax = b$, pero se agrega la restricción $x \geq 0$. ¿Cuándo existe una ***solución nonegativa*** de $Ax = b$?

En el capítulo 2, b estaba en cualquier parte del espacio columna. Ahora sólo se permiten combinaciones *nonegativas*, y las bs ya no llenan un subespacio. En vez de lo anterior, ocupan una región en forma de cono. Para n columnas en $\mathbf{R}^m$, el cono se convierte en una pirámide de extremo abierto. En la figura 8.4 se muestran cuatro vectores en $\mathbf{R}^2$, y A es de 2 por 4. Si b está en este cono, existe una solución nonegativa de $Ax = b$; en caso contrario, no.

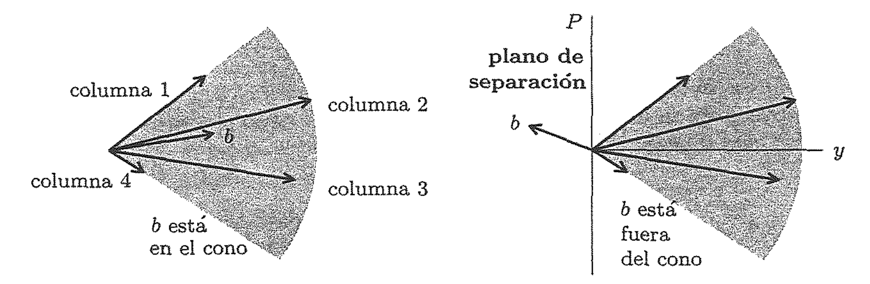


Figura 8.4 Cono de las combinaciones nonegativas de las columnas: $b = Ax$ con $x \geq 0$. Cuando b está fuera del cono, está separado por un hiperplano (perpendicular a y).

¿Cuál es la alternativa si b está fuera del cono? En la figura 8.4 también se muestra un "hiperplano de separación" que tiene, al vector b en un lado y a todo el cono en el otro lado. El plano consta de todos los vectores perpendiculares a un vector fijo y. El ángulo entre y y b es mayor que 90°, por lo que $yb < 0$. El ángulo entre y y toda columna de A es menor que 90°, de modo que $yA \geq 0$. Esta es la alternativa que se buscaba. Este *teorema del hiperplano de separación* es fundamental para la economía matemática.

8I $Ax = b$ tiene una solución **nonegativa, *o bien*,** existe una y con $yA \geq 0$ y $yb < 0$.

Ejemplo 1 Las combinaciones nonegativas de las columnas de $A = I$ llenan el cuadrante positivo $b \geq 0$. Para cualquier otra b, la alternativa debe cumplirse para alguna y:

No está en el cono Si $b = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$, entonces $y = \begin{bmatrix} 0 & 1 \end{bmatrix}$ proporciona $yI \geq 0$ pero $yb = -3$.

El eje x, perpendicular a $y = [0\ 1]$, separa b del cono = cuadrante.

A continuación se presenta un par de alternativas curiosas. Es imposible que un subespacio S y su complemento ortogonal $S^{\perp}$ contengan vectores positivos. El producto interno debe ser positivo, no cero. Sin embargo, S podría ser el eje x y $S^{\perp}$ podría ser el eje y, en cuyo caso contienen a los vectores "semipositivos" [1 0] y [0 1]. Esta alternativa ligeramente más débil funciona: *Ya sea que S contiene un vector positivo $x > 0$, o que $S^{\perp}$ contiene un $y \geq 0$ diferente de cero.* Cuando S y $S^{\perp}$ son rectas perpendiculares en el plano, una o la otra deben estar en el primer cuadrante. No puedo ver claramente esto en tres o cuatro dimensiones.

Para la programación lineal, las alternativas importantes provienen cuando las restricciones son desigualdades. ¿Cuándo ocurre que el conjunto factible es vacío (no x)?

8J $Ax \geq b$ tiene una solución $x \geq 0$ ***o bien***, existe una $y \leq 0$ con $yA \geq 0$ y $yb < 0$.

Demostración Las variables flojas $w = Ax - b$ cambian a $Ax \geq b$ en una ecuación. Se aplica 8I:

Primera alternativa $\quad \begin{bmatrix} A & -I \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} = b$ para algún $\begin{bmatrix} x \\ w \end{bmatrix} \geq 0.$

Segunda alternativa $\quad y\begin{bmatrix} A & -I \end{bmatrix} \geq \begin{bmatrix} 0 & 0 \end{bmatrix}$ para alguna y con $yb < 0$. ■

Este es el resultado que conduce a una "demostración no constructiva" del teorema de dualidad.

Conjunto de problemas 8.3

1. ¿Cuál es el dual del siguiente problema: Minimizar $x_1 + x_2$, sujeto a $x_1 \geq 0$, $x_2 \geq 0$, $2x_1 \geq 4$, $x_1 + 3x_2 \geq 11$? Encuentre la solución de este problema y su dual, y compruebe que el mínimo es igual al máximo.
2. ¿Cuál es el dual del siguiente problema: Maximizar y_2, sujeto a $y_1 \geq 0$, $y_2 \geq 0$, $y_1 + y_2 \leq 3$? Resuelva este problema y su dual.
3. Suponga que A es la matriz identidad (de modo que $m = n$), y que los vectores b y c son nonegativos. Explique por qué $x^* = b$ es óptimo en el problema del mínimo, en-

cuentre y^* en el problema del máximo, y compruebe que los dos valores son iguales. Si la primera componente de b es negativa, ¿cuáles son x^* y y^*?

4. Construya un ejemplo de 1 por 1 donde $Ax \geq b$, $x \geq 0$ es no factible, y el problema dual no esté acotado.

5. Empezando con la matriz de 2 por 2 $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, escoja b y c de modo que los dos conjuntos factibles $Ax \geq b, x \geq 0$ y $yA \leq c, y \geq 0$ sean vacíos.

6. Si todos los elementos de A, b, y c son positivos, demuestre que tanto el original como el dual son factibles.

7. Demuestre que $x = (1, 1, 1, 0)$ y $y = (1, 1, 0, 1)$ son factibles en el original y en el dual, con

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad c = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 3 \end{bmatrix}.$$

Luego, después de calcular cx y yb, explique cómo se sabe que son óptimos.

8. Verifique que los vectores en el ejercicio previo satisfacen las condiciones de flojedad complementarias en la ecuación (2), y encuentre la desigualdad floja tanto en el original como en el dual.

9. Suponga que $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, y $c = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Encuentre las x y y óptimas, y compruebe las condiciones de flojedad complementarias (así como $yb = cx$).

10. Si el problema original se restringe por ecuaciones en vez de por desigualdades —*Minimizar cx sujeto a $Ax = b$* y $x \geq 0$— luego, el requerimiento $y \geq 0$ se deja fuera del dual: *Maximizar yb sujeto a $yA \leq c$*. Demuestre que la desigualdad unilateral $yb \leq cx$ sigue cumpliéndose. ¿Por qué en la ecuación (1) se requería $y \geq 0$ pero aquí no es necesario? Esta dualidad débil puede completarse para dualidad total.

11. *a*) Sin utilizar el método simplex, minimice el costo $5x_1 + 3x_2 + 4x_3$, sujeto a $x_1 + x_2 + x_3 \geq 1, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$.
b) ¿Cuál es la forma del conjunto factible?
c) ¿Cuál es el problema dual, y cuál es su solución y?

12. Si el problema original tiene una solución única x^*, y luego c se modifica ligeramente, explique por qué x^* sigue siendo la solución óptima.

13. Escriba el dual del siguiente problema: Maximizar $x_1 + x_2 + x_3$ sujeto a $2x_1 + x_2 \leq 4$, $x_3 \leq 6$. ¿Cuáles son las x^* y y^* óptimas (¡en caso de existir!)?

14. Si $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, describa el cono de combinaciones nonegativas de las columnas. Si b está dentro del cono, por ejemplo $b = (3, 2)$, ¿cuál es el vector factible x? Si b está fuera, por ejemplo $b = (0, 1)$, ¿qué vector y sigue satisfaciendo la alternativa?

15. En tres dimensiones, ¿es posible encontrar un conjunto de seis vectores cuyo cono de combinaciones nonegativas llene todo el espacio? ¿Qué puede decir sobre cuatro vectores?

16. Use 8H para demostrar que la siguiente ecuación no tiene solución, ya que la alternativa se cumple:

$$\begin{bmatrix} 2 & 2 \\ 4 & 4 \end{bmatrix} x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

17. Use 8I para demostrar que no hay solución $x \geq 0$ (la alternativa se cumple):

$$\begin{bmatrix} 1 & 3 & -5 \\ 1 & -4 & -7 \end{bmatrix} x = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

18. Demuestre que las alternativas en 8J ($Ax \geq b, x \geq 0, yA \geq 0, yb < 0, y \leq 0$) no pueden cumplirse. *Sugerencia*: yAx.

8.4 MODELOS DE REDES

Algunos problemas lineales tienen una estructura que llega muy rápido a su solución. Las matrices de banda tienen elementos diferentes de cero cerca de la diagonal principal, y es fácil resolver $Ax = b$. En programación lineal, se tiene interés en la clase especial para la cual A es una ***matriz de incidencia***. Sus elementos son -1 o $+1$, o (en su mayoría) cero, y los pasos de pivoteo sólo implican sumas y restas. Es posible resolver problemas mucho más grandes que los de costumbre.

Las redes se presentan en todo tipo de aplicaciones. La circulación en un crucero satisface las leyes de la corriente de Kirchhoff: el flujo que entra es igual al flujo que sale. Para gas y petróleo, la programación de redes ha diseñado sistemas de conducción que son millones de dólares más baratos que los diseños intuitivos (no optimizados). ¡La programación de pilotos, tripulaciones y aviones se ha convertido en un problema importante en matemáticas aplicadas! Incluso se resuelve el ***problema del matrimonio***: maximizar el número de matrimonios cuando las novias están vetadas. Este problema puede no ser real, pero es uno que resuelve la programación de redes.

El problema en la figura 8.5 consiste en ***maximizar el flujo de la fuente al sumidero***. Los flujos no pueden exceder las capacidades indicadas en las aristas, y las direcciones proporcionadas por las flechas no pueden invertirse. El flujo sobre las dos aristas hacia el sumidero no puede exceder $6 + 1 = 7$. ¿Es posible alcanzar este total de 7? ¿Cuál es el ***flujo máximo*** de izquierda a derecha?

Las incógnitas son los flujos x_{ij} del nodo i al nodo j. Las restricciones de capacidad son $x_{ij} \leq c_{ij}$. Los flujos son nonegativos: $x_{ij} \geq 0$ en dirección de las flechas. Al maximizar el flujo de regreso x_{61} (la recta punteada), se maximiza el flujo total hacia el sumidero.

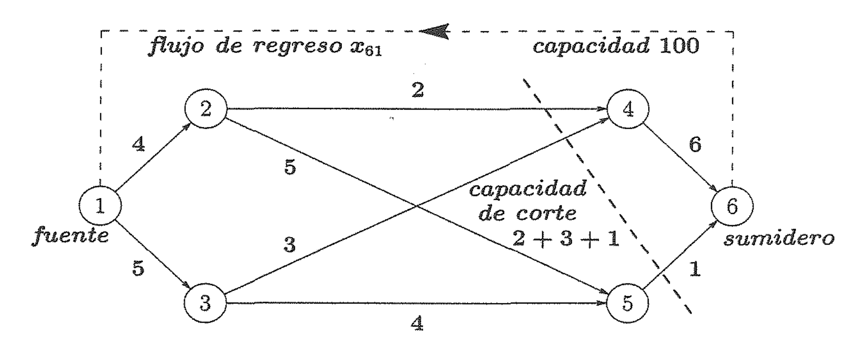


Figura 8.5 Red con 6 nodos con aristas de capacidad: el problema del flujo máximo.

Aún se escuchará hablar de otra restricción. Se trata de la "ley de conservación": *el flujo que entra en cada nodo es igual al flujo que sale de cada nodo*. Esta es la ley de la corriente de Kirchhoff:

$$\textbf{Ley de la corriente} \qquad \sum_i x_{ij} - \sum_k x_{jk} = 0 \quad \text{para} \quad j = 1, 2, \ldots, 6. \tag{1}$$

Los flujos x_{ij} entran al nodo j provenientes de nodos previos i. Los flujos x_{jk} salen del nodo j a nodos posteriores k. El balance de la ecuación (1) puede escribirse como $Ax = 0$, donde A es una *matriz de incidencia nodo-arista* (la traspuesta de la sección 2.5). A tiene una flecha por cada nodo y una columna $+1$, -1 para cada arista:

$$\textbf{Matriz de incidencia} \quad A = \begin{bmatrix} 1 & 1 & & & & & & & -1 \\ -1 & & 1 & 1 & & & & & \\ & -1 & & & 1 & 1 & & & \\ & & -1 & & -1 & & 1 & & \\ & & & -1 & & -1 & & 1 & \\ & & & & & & -1 & -1 & 1 \end{bmatrix} \begin{matrix} \textbf{nodo } 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix}$$

$$\textbf{arista} \quad 12 \quad 13 \quad 24 \quad 25 \quad 34 \quad 35 \quad 46 \quad 56 \quad 61$$

$$\textbf{Flujo máximo} \qquad \textbf{Maximizar } x_{61} \textbf{ sujeto a } Ax = 0 \textbf{ y } 0 \leq x_{ij} \leq c_{ij}.$$

Un flujo de 2 puede ir en la trayectoria 1-2-4-6-1. Un flujo de 3 puede ir a lo largo de 1-3-4-6-1. Un flujo adicional de 1 puede seguir la trayectoria más corta 1-3-5-6-1. El total es 6, y *nada más es posible*. ¿Cómo se demuestra que el flujo maximal es 6 y no 7?

El método al tanteo es convincente, pero las matemáticas son concluyentes: La clave es encontrar un ***corte*** en la red, a través del cual todas las capacidades se llenen. Este corte separa los nodos 5 y 6 de los demás nodos. Las aristas que van directamente a través del corte tienen una capacidad total de $2 + 3 + 1 = 6$, ¡y nada más puede cruzar! La dualidad débil establece que cada corte constituye una cota para el flujo total, y la dualidad total establece que el corte de menor capacidad (el *corte minimal*) se llena por el flujo maximal.

8K ***Teorema del flujo máx-corte mín*** El flujo maximal en una red es igual a la capacidad total a través del corte minimal.

Un "corte" separa los nodos en dos grupos S y T (fuente en S y sumidero en T). Su capacidad es la suma de las capacidades de todas las aristas que cruzan desde el corte (de S a T). Varios cortes podrían tener la misma capacidad. Ciertamente, el flujo total nunca puede ser mayor que la capacidad total a través del corte minimal. El problema, aquí y en todo lo que concierne a la dualidad, es demostrar que la igualdad se alcanza gracias al flujo y el corte idóneos.

Demostración de que flujo máx = corte mín Suponga que un flujo es maximal. Algunos nodos aún podrían alcanzarse desde la fuente mediante flujo adicional, sin exceder ninguna de las capacidades. Estos nodos van con la fuente hacia el conjunto S. El sumidero debe estar en el conjunto restante T, o bien, ¡hubieran recibido más flujo! Cada arista a través del corte debe llenarse, o flujo adicional podría avanzar más lejos hacia el nodo en T. Así, el flujo maximal llena este corte a su capacidad y se ha alcanzado la igualdad. ■

Lo anterior sugiere una forma de construir el flujo maximal: Comprobar si alguna trayectoria posee una capacidad no utilizada. En caso de encontrar una, agregar flujo a lo largo de la "trayectoria aumentada". Luego, calcular las capacidades restantes y decidir si el sumidero está cortado respecto a la fuente o si es posible añadir flujo adicional. Si cada nodo en S se identifica con el nodo previo del que pudo provenir el flujo, entonces es posible realizar un rastreo para encontrar la trayectoria del flujo adicional.

El problema del matrimonio

Suponga que hay cuatro mujeres y cuatro hombres. Algunas de estas 16 parejas son compatibles y otras no. ¿Cuándo es posible encontrar un ***emparejamiento completo***, en la que todos estén casados? Si el álgebra lineal es capaz de trabajar en el espacio de dimensión 20, ciertamente puede manejar el problema trivial del matrimonio.

Hay dos formas de presentar el problema: en una matriz o en una gráfica. La matriz contiene $a_{ij} = 0$ si la i-ésima mujer y el j-ésimo hombre no son compatibles, y $a_{ij} = 1$ si están intentando ser compatibles. Así, el renglón i proporciona las opciones de la i-ésima mujer, y la columna j corresponde al j-ésimo hombre:

$$\textbf{Matriz de compatibilidad} \qquad A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \text{tiene 6 pares compatibles}$$

La gráfica de la izquierda en la figura 8.6 muestra dos matrimonios posibles. Ignorando la fuente s y el sumidero t, tiene cuatro mujeres en la izquierda y cuatro hombres en la derecha. *Las aristas corresponden a los* 1s *en la matriz*, y las capacidades son 1 matrimonio. Entre la primera mujer y el cuarto hombre no hay aristas, ya que la matriz tiene $a_{14} = 0$.

Podría parecer que no es posible alcanzar el nodo M_2 por más flujo, ¡pero éste no es el caso! El flujo adicional a la derecha va de regreso para cancelar un matrimonio existente. Este flujo adicional realiza tres matrimonios, lo que es maximal. El corte minimal lo cruzan 3 aristas.

Un emparejamiento completo (en caso de ser posible) es un conjunto que tiene cuatro 1s en la matriz. Éstos pueden provenir de cuatro renglones distintos y cuatro columnas diferentes, ya que no se permite la bigamia. Es como encontrar una *matriz de permutación* dentro de los elementos de A diferentes de cero. En la gráfica, esto significa cuatro aristas sin ningún nodo en común. El flujo maximal es menor que 4 exactamente cuando un emparejamiento completo es imposible.

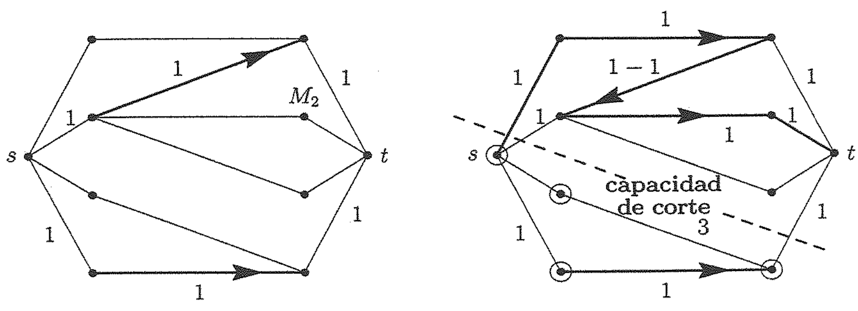


Figura 8.6 Dos matrimonios a la izquierda, tres (máximo) a la derecha. El tercero es creado agregando dos nuevos matrimonios y un divorcio (flujo hacia atrás).

En nuestro ejemplo, el flujo maximal es 3, no 4. Se permiten los matrimonios 1-1, 2-2, 4-4 (así como varios otros conjuntos de tres matrimonios), aunque no hay forma de lograr cuatro. El corte minimal a la derecha separa a las dos mujeres en el fondo de los tres hombres en la parte superior. Las dos mujeres sólo tienen un hombre restante para elegir, lo cual no es suficiente. La capacidad a través del corte es sólo de 3.

Siempre que hay un subconjunto de k mujeres que deben decidir entre un conjunto menor que k hombres, un emparejamiento completo es imposible.

La prueba es decisiva. La misma imposibilidad puede plantearse en varias formas:

1. **(Para ajedrez)** Resulta imposible poner cuatro peones en escaques que tienen 1s en A, de modo que bajo enroque no es posible, puede comer a cualquier otro enroque.
2. **(Para matrices de matrimonio)** Los 1s en la matriz pueden cubrirse con tres líneas horizontales o verticales. Esto es igual al número de matrimonios.
3. **(Para álgebra lineal)** Toda matriz con los mismos ceros que A es singular.

Recuerde que el determinante es una suma de $4! = 24$ términos. En cada término se utilizan los cuatro renglones y columnas. Los ceros en A hacen cero a todos los 24 términos.

¡Un bloque de ceros impide un emparejamiento completo! La submatriz de 2 por 3 en los renglones 3 y 4 y en las columnas 1, 2, 3 de A es completamente cero. La regla general para una matriz de n por n es que ***un bloque p por q de ceros impide un emparejamiento si $p + q > n$.*** Aquí las mujeres 3, 4 sólo podrían casarse con 4 hombres. Si p mujeres pueden casarte con $n - q$ y $p > n - q$ hombres (lo cual es igual al bloque cero con $p + q > n$), de modo que un emparejamiento completo es imposible.

El problema matemático es demostrar lo siguiente: ***si a todo conjunto de p mujeres le agrada por lo menos p hombres, un emparejamiento completo es imposible. Esta es la condición de Hall.*** Ningún bloque de ceros es demasiado grande. Cada mujer debe agradarle por lo menos a un hombre; dadas dos mujeres, una de ellas debe agradarle por lo menos a dos hombres, y así sucesivamente, hasta $p = n$.

8L Un emparejamiento completo es posible si (y sólo si) se cumple la condición de Hall.

La demostración es más sencilla si las capacidades son n, en vez de 1, sobre todas las aristas a través de la parte de en medio. Las capacidades que salen de la fuente y se dirigen hacia el sumidero siguen siendo 1. Si el flujo maximal es n, entonces las aristas que van de la fuente al sumidero están llenas, y el flujo produce n matrimonios. Cuando un emparejamiento completo es imposible y el flujo máximo está por abajo de n, algún corte debe ser responsable.

Ese corte mantiene a la capacidad por abajo de n, de modo que las aristas de en medio la cruzan. Suponga que p nodos a la izquierda y r nodos a la derecha están en el conjunto S con la fuente. La capacidad a través de ese corte está a $n - p$ de la fuente a las mujeres restantes, y a r de estos hombres al sumidero. Debido a que la capacidad de corte está por debajo de n, *las p mujeres sólo agradan a los r hombres* y a ningún otro. Sin embargo, la capacidad $n - p + r$ está por abajo de n exactamente cuando $p > r$, y la condición de Hall fracasa.

Árboles generadores y el algoritmo avaro

Uno de los modelos fundamentales de redes es el ***problema de la ruta mínima***, donde las aristas tienen *longitudes* en vez de capacidades. Se busca la ruta más corta de la fuente al sumidero. Si las aristas son líneas telefónicas y las longitudes son tiempos de retraso, se está encontrando la ruta más rápida para una llamada. Si los nodos son computadoras, se está buscando el protocolo perfecto para pasar el mensaje.

Un problema estrechamente relacionado encuentra el ***árbol generador más corto***: un conjunto de $n - 1$ aristas que unen todos los nodos de la red. En vez de ir rápidamente de una fuente a un sumidero, ahora se está minimizando el costo de conectar *todos* los nodos. No hay ciclos porque el costo de cerrar un ciclo es innecesario. *Un árbol generador conecta los nodos sin ciclos*, y se desea encontrar el más corto. A continuación se presenta un algoritmo posible:

1. *Se empieza en cualquier nodo s y se repite el paso siguiente:*
   Sumar la arista más corta que une el árbol actual a un nuevo nodo.

En la figura 8.7, las longitudes de las aristas se observan en el orden 1, 2, 7, 4, 3, 6. El último paso omite la arista de longitud 5, que cierra un ciclo. La longitud total es 23, pero, ¿es mínimo? La arista de longitud 7 se aceptó muy fácil, y el segundo algoritmo proporciona una longitud más larga.

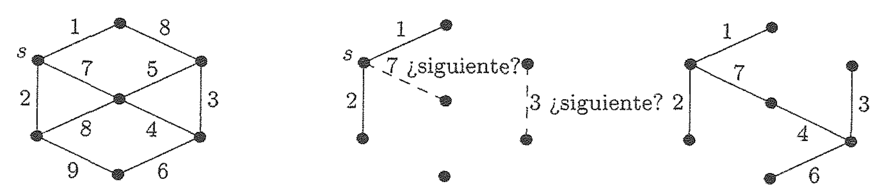


Figura 8.7 Red y árbol generador más corto de longitud 23.

2. *Aceptar las aristas en orden creciente de longitud, rechazando las aristas que completan un ciclo.*

Ahora las aristas se presentan en el orden 1, 2, 3, 4, 6 (rechazando 5 de nuevo), y 7. Son las mismas aristas, aunque esto no ocurre siempre. Su longitud total es la misma, lo cual *ocurre* siempre. ***El problema del árbol generador es excepcional, porque puede resolverse en un paso.***

En el lenguaje de la programación lineal, primero se está encontrando el vértice óptimo. El problema del árbol generador se resuelve como en la sustitución hacia atrás, *sin pasos falsos*. Este método general se denomina ***algoritmo avaro***. A continuación se presenta otra idea codiciosa:

3. *Construir árboles a partir de todos los n nodos, repitiendo el paso siguiente:*
   Seleccionar cualquier árbol y sumar la arista de longitud mínima que sale de ese árbol.

Los pasos dependen del orden de selección de los árboles. Para permanecer en el mismo árbol es el algoritmo **1**. Tomar las longitudes en orden es el algoritmo **2**. Barrer a través de todos los árboles a la vez es un nuevo algoritmo. Se oye muy fácil, pero para un gran problema la estructura de los datos se vuelve crítica. Con mil nodos, podría haber cerca de un millón de aristas, y no es recomendable repasar esta lista mil veces.

Modelos de redes adicionales

Los siguientes problemas están relacionados para hacer coincidir y son casi tan fáciles:

1. El ***problema de la asignación óptima***: a_{ij} mide el valor del aspirante i al puesto j. Asignar puestos para maximizar el valor total —la suma de los a_{ij} sobre los puestos asignados. (Si todos los a_{ij} son 0 o 1, este es el problema del matrimonio.)

2. El *problema de transporte*: Dados suministros en n puntos y demandas en n mercados, escoger embarques x_{ij} de los proveedores a los mercados que minimicen el costo total $\sum C_{ij}x_{ij}$. (Si todos los suministros y las demandas son 1, este es el problema de la asignación óptima: enviar una persona a cada puesto.)

3. ***Flujo de costo mínimo***: Ahora las rutas tienen capacidades c_{ij}, así como costos C_{ij}, mezclando el problema de flujo máximo con el problema de transporte. ¿Cuál es el flujo más barato, sujeto a restricciones de capacidad?

Una parte fascinante de este tema es el desarrollo de algoritmos. En vez de una demostración teórica de la dualidad, se usa *primera búsqueda de amplitud* o *primera búsqueda de profundidad* para encontrar la asignación óptima o el flujo más barato. Es como el método simplex, ya que empieza con un flujo factible (un vértice) y suma un nuevo flujo (el desplazamiento hacia el siguiente vértice). Los algoritmos son especiales porque los problemas de redes implican matrices de incidencia.

La técnica de ***programación dinámica*** se apoya en un concepto simple: Si una ruta de una fuente a un sumidero es óptima, entonces *cada parte de la ruta debe ser óptima*. La solución es construir hacia atrás a partir del sumidero, en un proceso de decisión de etapas múltiples. En cada etapa, la distancia al sumidero es el mínimo de una nueva distancia más una distancia anterior:

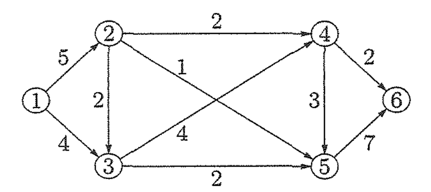
Ecuación de Bellman distancia $\boldsymbol{x}$-$\boldsymbol{t}$ = mínimo sobre y de (distancias $\boldsymbol{x}$-$\boldsymbol{y}$ + $\boldsymbol{y}$-$\boldsymbol{t}$).

Me hubiese gustado tener más espacio sobre redes. Son simples y hermosas.

Conjunto de problemas 8.4

1. En la figura 8.5, sume 3 a cada capacidad. Por inspección, encuentre el flujo máximo y el corte mínimo.

2. Encuentre un flujo máximo y el corte mínimo para la siguiente red:



3. Si es posible incrementar la capacidad de cualquier tubo en la red de la figura anterior, ¿qué cambio provocaría el mayor incremento en el flujo máximo?

4. Trace una red de cinco nodos con capacidad $|i - j|$ entre el nodo i y el nodo j. Encuentre el máximo flujo posible del nodo 1 al nodo 4.

5. En una gráfica, el número máximo de rutas de s a t sin aristas comunes es igual al número mínimo de aristas cuya eliminación desconecta s de t. Relacione esto con el teorema del flujo máximo-corte mínimo.

6. Encuentre un conjunto máximo de matrimonios (un emparejamiento completo, de ser posible) para

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Trace la red para B, con líneas más gruesas en las aristas de los emparejamientos.

7. Para la matriz A del problema 6, ¿qué renglones violan la condición de Hall al tener todos sus 1s en muy pocas columnas? ¿Qué submatriz de ceros de p por q cumple $p + q > n$?

8. ¿Cuántas rectas (horizontales y verticales) se requieren para cubrir todos los unos en la matriz A del problema 6? Para cualquier matriz, explique por qué la dualidad débil es cierta: si k matrimonios son posibles, entonces se requieren por lo menos k rectas para cubrir todos los 1s.

9. *a*) Suponga que cada renglón y cada columna contienen exactamenten dos 1s. Demuestre que un emparejamiento completo es posible. (Demuestre que no es posible cubrir los 1s con menos de n rectas.)
   b) Encuentre un ejemplo con dos *o más* 1s en cada renglón y cada columna, para el cual un emparejamiento completo es imposible.

10. Si una matriz de 7 por 7 tiene 15 1s, demuestre que permite por lo menos tres matrimonios.

11. Para conjuntos *infinitos*, un emparejamiento completo puede ser imposible incluso si se cumple la condición de Hall. Si todo el primer renglón es de 1s y entonces cada $a_{ii-1} = 1$, demuestre que p renglones cualesquiera contienen 1s en por lo menos p columnas; y ni así se tiene un emparejamiento completo.

12. Si la figura 8.5 muestra longitudes en vez de capacidades, encuentre la ruta más corta de s a t, y un árbol generador mínimo.

13. Aplique los algoritmos 1 y 2 para encontrar un árbol generador mínimo para la red del problema 2.

14. *a*) ¿Por qué el algoritmo avaro es válido para el problema del árbol generador?
    b) Demuestre con un ejemplo que el algoritmo avaro podría fracasar en encontrar la ruta más corta de s a t, empezando con la arista más corta.

15. Si A es la matriz de 5 por 5 con 1s justo arriba y abajo de la diagonal principal, encuentre
    a) Un conjunto de renglones con 1s en muy pocas columnas.
    b) Un conjunto de columnas con 1s en muy pocos renglones.
    c) Una submatriz de ceros de p por q con $p + q > 5$.
    d) Cuatro rectas que cubran a todos los 1s.

16. El problema de flujo máximo tiene variables flojas $w_{ij} = c_{ij} - x_{ij}$ para la diferencia entre capacidades y flujos. Formule el problema de la figura 8.5 como un programa lineal.

8.5 TEORÍA DE JUEGOS

La mejor manera de explicar un ***juego de suma-cero de dos personas*** es mediante un ejemplo. Hay dos jugadores X y Y, y las reglas son las mismas en cada turno:

X tiene una mano o dos, así como Y. Si toman la misma decisión, Y gana \$10. Si toman decisiones opuestas, X gana \$10 por una mano y \$20 por dos:

Matriz de pagos (pagos a X)

$$A = \begin{bmatrix} -10 & 20 \\ 10 & -10 \end{bmatrix} \begin{matrix} \text{una mano por } Y \\ \text{dos manos por } Y \end{matrix}$$

una mano por X — dos manos por X

Si X hace lo mismo cada vez, Y lo imitará y ganará. De manera semejante, Y no puede adherirse a una sola estrategia, o X hará lo contrario. Ambos jugadores deben usar una ***estrategia mezclada***, y la elección en cada turno debe ser independiente de los turnos previos. Si hay algún patrón histórico, el oponente puede aprovecharlo. Incluso la estrategia "quedarse con la misma opción hasta perder" es evidentemente fatal. Luego de suficientes juegos, el oponente debe saber exactamente qué esperar.

En una estrategia mezclada, X puede poner una mano con una frecuencia x_1 y ambas manos con frecuencia $x_2 = 1 - x_1$. En cada turno, esta decisión es aleatoria. De manera semejante, Y puede escoger probabilidades y_1 y $y_2 = 1 - y_1$. Ninguna de estas probabilidades debe ser 0 o 1; en caso contrario, el oponente ajusta su estrategia y gana. Si son iguales a $\frac{1}{2}$, Y perdería \$20 muy a menudo. (Podría perder \$20 la cuarta parte del tiempo, \$10 otra cuarta parte del tiempo, y ganar \$10 la mitad del tiempo, lo cual supone una pérdida media de \$2.50. Esto es más de lo necesario.) Pero mientras Y se mueve más hacia una estrategia pura de dos manos, más se moverá X hacia una mano.

El problema fundamental es *encontrar las mejores estrategias mezcladas*. ¿X puede escoger probabilidades x_1 y x_2 que presenten a Y sin razón para moverse en su propia estrategia (y viceversa)? Entonces el pago medio habrá alcanzado un ***punto silla***: es máximo en tanto X esté preocupado, y es mínimo en tanto Y esté preocupado. Encontrar este punto silla es resolver el juego.

X combina las dos columnas con pesos x_1 y $1 - x_1$ para producir una nueva columna "mezclada". Los pesos $\frac{3}{5}$ y $\frac{2}{5}$ deben producir esta columna:

Columna mezclada
$$\frac{3}{5}\begin{bmatrix} -10 \\ 10 \end{bmatrix} + \frac{2}{5}\begin{bmatrix} 20 \\ -10 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

***Contra esta estrategia mezclada, Y siempre pierde* \$2.** ¡Esto no significa que todas las estrategias son óptimas para Y! Si Y es flojo y se queda con una mano, X cambiará y empezará a ganar \$20. Luego Y cambiará, y luego también lo hará X. Finalmente, como se supone que ambos son inteligentes, se quedarán en estrategias mezcladas óptimas. Y combinará los *renglones* con los pesos y_1 y $1 - y_1$, tratando de producir un nuevo renglón que sea lo más pequeño posible:

Renglón mezclado $y_1\begin{bmatrix} -10 & 20 \end{bmatrix} + (1 - y_1)\begin{bmatrix} 10 & -10 \end{bmatrix} = \begin{bmatrix} 10 - 20y_1 & -10 + 30y_1 \end{bmatrix}.$

La mezcla correcta hace iguales a las dos componentes, en $y_1 = \frac{2}{5}$. Luego, hace iguales a 2 a ambas componentes; el renglón mezclado se vuelve [2 2]. ***Con esta estrategia Y no puede perder más de* \$2.**

Y ha minimizado la pérdida máxima, y este *minimax* coincide con el *maximin* encontrado por X. El *valor del juego* es minimax = maximin = \$2.

¡La mezcla óptima de renglones puede no siempre tener elementos iguales! Suponga que a X se permite una tercera estrategia de mantener hasta tres manos para ganar \$60 cuando Y pone una mano y \$80 cuando Y pone hasta dos. La matriz de pagos se convierte en

$$A = \begin{bmatrix} -10 & 20 & 60 \\ 10 & -10 & 80 \end{bmatrix}.$$

X escogerá la estrategia de las tres manos (columna 3) cada vez, y ganará por lo menos \$60. Al mismo tiempo, Y siempre escoge el primer renglón; la pérdida máxima es \$60. Sigue teniéndose maximin = minimax = \$60, pero el punto silla está en el vértice.

En la mezcla óptima de renglones de Ys, que era puramente el renglón 1, \$60 aparece sólo en la columna realmente utilizada por X. En la mezcla óptima de columnas de Xs, que era la columna 3, \$60 aparece en el renglón que entra en la mejor estrategia de Ys. Esta regla corresponde exactamente a la *condición floja complementaria* de la programación lineal.

Juegos matriciales

El "juego matricial de m por n" más general, es exactamente como nuestro ejemplo. X tiene n posibles movimientos (columnas de A). Y escoge de los m renglones. El elemento a_{ij} está en la matriz de pagos cuando X escoge la columna j y Y escoge el renglón i. Un elemento negativo significa un pago a Y. Este es un ***juego de suma-cero***. Lo que pierde un jugador lo gana el otro.

X es libre de escoger cualquier estrategia mezclada $x = (x_1, \ldots, x_n)$. Estos x_i proporcionan las frecuencias para las n columnas, cuya suma es 1. En cada turno, X utiliza un mecanismo aleatorio para producir la estrategia i con frecuencia x_i. Y escoge un vector $y = (y_1, \ldots, y_m)$, también con $y_i \geq 0$ y $\Sigma y_i = 1$, que proporciona las frecuencias para seleccionar los renglones.

Una simple jugada del juego es aleatoria. En promedio, la combinación de la columna j para X y del renglón i para Y aparece con probabilidad $x_j y_i$. Cuando se presenta, el pago es a_{ij}. El pago esperado para X a partir de esta combinación es $\sum\sum a_{ij}x_j y_i = yAx$:

$$yAx = \begin{bmatrix} y_1 & \cdots & y_m \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = a_{11}x_1y_1 + \cdots + a_{mn}x_ny_m = \textbf{Pago medio.}$$

Este es el pago yAx que X quiere maximizar y Y quiere minimizar.

Ejemplo 1 Suponga que A es la matriz identidad de n por n, $A = I$. El pago esperado se vuelve $yIx = x_1y_1 + \cdots + x_ny_n$. X espera encontrar la misma opción que Y, ganar $a_{ii} = \$1$. Y está intentando evadir a X, pagar $a_{ij} = \$0$. Si X escoge cualquier columna más a menudo que otra, Y puede escapar más a menudo. *La mezcla óptima es* $x^* = (1/n, 1/n, \ldots, 1/n)$. De igual manera, Y no puede preferir a ningún renglón, ya que la mezcla óptima es $y^* = (1/n, 1/n, \ldots, 1/n)$. La probabilidad de que ambos escojan la estrategia i es $(1/n)^2$, y la sumatoria sobre i es el pago esperado a X. El valor total del juego es n veces $(1/n)^2$, o $1/n$:

$$y^*Ax^* = \begin{bmatrix} 1/n & \cdots & 1/n \end{bmatrix} \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} \begin{bmatrix} 1/n \\ \vdots \\ 1/n \end{bmatrix} = \left(\frac{1}{n}\right)^2 + \cdots + \left(\frac{1}{n}\right)^2 = \frac{1}{n}.$$

Cuando n crece, Y tiene mejores posibilidades de escapar. El valor $1/n$ baja.

La matriz simétrica $A = I$ no hizo justo el juego. Una *matriz simétrica sesgada*, $A^T = -A$, *significa un juego completamente justo*. Así, una elección de la estrategia j por X e i por Y gana a_{ij} para X, y la elección de j por Y e i por X gana la misma cantidad para Y (porque $a_{ji} = -a_{ij}$). Las estrategias óptimas x^* y y^* deben ser las mismas, y el pago esperado debe ser $y^*Ax^* = 0$. El valor del juego, cuando $A^T = -A$, es cero. Aunque aún es necesario encontrar la estrategia.

Ejemplo 2

$$\textbf{Juego justo} \qquad A = \begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix}.$$

En palabras, ambos X y Y escogen un número entre 1 y 3. La elección más pequeña gana \$1. (Si X escoge 2 y Y escoge 3, el pago es $a_{32} = \$1$; si escogen el mismo número, se está en la diagonal y nadie gana.) Ningún jugador puede escoger una estrategia que implique 2 o 3. Las estrategias puras $x^* = y^* = (1, 0, 0)$ son óptimas: ambos jugadores escogen 1 cada vez. El valor es $y^*Ax^* = a_{11} = 0$.

La matriz que deja sin cambio a todas las decisiones tiene mn elementos iguales, por ejemplo α. Esto significa simplemente que X gana una cantidad adicional α en cada turno. El valor del juego se incrementa por α, pero no hay razón para cambiar a x^* y y^*.

El teorema minimax

Póngase en el lugar de X, quien escoge la estrategia mezclada $x = (x_1, \ldots, x_n)$. Y terminará por reconocer esa estrategia y escogerá y para *minimizar* el pago yAx. Un jugador inteligente X elegirá x^* para ***maximizar este mínimo***:

$$\textbf{X gana por lo menos} \qquad \min_y yAx^* = \max_x \min_y yAx. \qquad (1)$$

El jugador Y hace lo contrario. Para cualquier estrategia elegida y, X *maximizará* yAx. En consecuencia, Y escogerá la mezcla y^* que ***minimiza este máximo***:

$$\textbf{Y pierde no más de} \qquad \max_x y^*Ax = \min_y \max_x yAx. \qquad (2)$$

Espero que el lector se dé cuenta cuál será el resultado clave, si es cierto. Se quiere que la cantidad en la ecuación (1) que se garantiza que X ganará sea igual a la cantidad en la ecuación (2) que Y debe estar satisfecho de perder. Así se resolverá el juego: X puede ganar sólo moviéndose a partir de x^* y Y puede perder sólo moviéndose a partir de y^*. La existencia de este punto silla fue demostrada por Von Neumann.

8M Para cualquier matriz A, el minimax sobre todas las estrategias es igual al máximo:

$$\textbf{Teorema minimax} \qquad \max_x \min_y yAx = \min_y \max_x yAx = \text{valor del juego.} \qquad (3)$$

Si el máximo en la izquierda se alcanza en x^*, y el mínimo en la derecha se alcanza en y^*, este es un punto silla a partir del que nadie quiere moverse:

$$y^*Ax \le y^*Ax^* \le yAx^* \quad \text{para toda } x \text{ y } y \qquad (4)$$

En este punto silla, x^* es por lo menos tan bueno como cualquier otro x (ya que $y^*Ax \le y^*Ax^*$). Y el segundo jugador Y sólo puede pagar más si abandona y^*.

Así como en la teoría de dualidad, *maximin* $\le$ *minimax* es fácil. En la ecuación (1) se combina la definición de x^*, y en la ecuación (2) se combina la definición de y^*:

$$\max_x \min_y yAx = \min_y yAx^* \le y^*Ax^* \le \max_x y^*Ax = \min_y \max_x yAx. \qquad (5)$$

Esto sólo afirma que si es posible garantizar que X gane por lo menos α, y que Y pierda no más de β, entonces $\alpha \le \beta$. El logro de Von Neumann fue demostrar que $\alpha = \beta$. El teorema minimax significa que la igualdad debe cumplirse en toda la ecuación (5).

Para nosotros, lo extraordinario sobre la demostración es que *utiliza exactamente los mismos procedimientos matemáticos que la teoría de la programación lineal*. X y Y desempeñan roles "duales". Ambos escogen estrategias del "conjunto factible" de vectores de probabilidad: $x_i \ge 0, \sum x_i = 1, y_i \ge 0, \sum y_i = 1$. Lo emocionante es que incluso Von Neumann no reconoció de inmediato ambas teorías como la misma. (Demostró el minimax en 1928, la programación lineal comenzó antes de 1947, y Gale, Kuhn, y Tucker publicaron la primera demostración de la dualidad en 1951, ¡con base en notas de Von Neumann!) Estamos invirtiendo la historia para deducir el teorema minimax a partir de la dualidad.

En breve, el teorema minimax puede demostrarse como sigue. Sean b el vector columna de m 1s, y c el vector renglón de n 1s. Estos programas lineales son duales:

minimizar cx,	maximizar yb,
sujeto a $Ax \ge b, x \ge 0$	sujeto a $yA \le c, y \ge 0$.

Para tener la certeza de que ambos problemas son factibles, a todos los elementos de A se suma un gran número α. Esto no puede afectar las estrategias óptimas, ya que cada pago asciende por α. Para la matriz resultante, que sigue denotándose por A, $y = 0$ es factible en el dual y cualquier x grande es factible en el original.

El teorema de dualidad de programación lineal garantiza x^* y y^* óptimos con $cx^* = y^*b$. Debido a los 1s que hay en b y en c, esto significa que $\sum x_i^* = \sum y_i^* = S$. La división entre S cambia las sumas a 1, y *las estrategias mezcladas resultantes* x^*/S *y* y^*/S *son óptimas*. Para cualesquiera otras estrategias x y y,

$$Ax^* \ge b \quad \text{implica} \quad yAx^* \ge yb = 1 \quad \text{y} \quad y^*A \le c \quad \text{implica} \quad y^*Ax \le cx = 1.$$

La cuestión más importante es que $y^*Ax \le 1 \le yAx^*$. Al dividir entre S, esto indica que el jugador X no puede ganar más que $1/S$ contra la estrategia y^*/S, y que el jugador X no puede perder más que $1/S$ contra x^*/S. Estas estrategias proporcionan maximin = minimax = $1/S$.

Juegos verdaderos

Con esto se completa la teoría, aunque queda una pregunta natural: ¿Cuáles juegos normales son equivalentes a los "juegos matriciales"? ***El ajedrez, el bridge y el póquer, ¿se ajustan a la teoría de Von Neumann?***

Considero que el ajedrez no se ajusta muy bien, por dos razones. Una estrategia para las negras debe incluir una decisión de cómo contestar a la primera jugada de las blancas, a la segunda jugada y así sucesivamente hasta el final del juego. X y Y tienen millones de estrategias puras. No veo mucho sobre el papel del azar. Si las blancas pueden encontrar una estrategia ganadora, o si las negras pueden encontrar una estrategia inspirada —ninguna de estas posibilidades se ha descubierto— eso podría efectivamente terminar el juego del ajedrez. Puede jugarse como tic-tac-toe (gato), aunque la emoción se perdería.

El bridge contiene algo decepcionante, como en un impasse. Se considera como un juego matricial, aunque de nuevo m y n son fantásticamente grandes. Quizá sea posible analizar parte por separado del bridge para encontrar una estrategia óptima. Lo mismo es válido para el béisbol, donde el pítcher y el bateador tratan de adivinar la elección del otro acerca del lanzamiento. (O bien el cátcher intenta adivinar cuándo el corredor se robará una base. Una "bola" cada vez otorga al bateador "base por bolas", de modo que debe haber una frecuencia óptima —dependiendo del corredor en base y de la situación.) De nuevo, una parte del juego puede aislarse para su análisis.

Por otro lado, el *blackjack no es un juego matricial* (en un casino) porque sigue reglas fijas. Mi amigo Ed Thorp encontró una estrategia ganadora mediante el conteo de cartas altas, obligando a barajar más y más naipes en Las Vegas. No había ningún elemento de azar, y tampoco ninguna estrategia x^*. El éxito editorial *Bringing Down the House* afirma cómo estudiantes del MIT ganaron bastante dinero (mientras no estaban haciendo sus deberes).

También está el ***dilema del prisionero***, en el que a dos cómplices se les ofrece por separado el mismo trato: confiesa y eres libre, en el supuesto de que tu cómplice no confiese (así, el cómplice obtiene 10 años de prisión). Si ambos confiesan, cada uno es condenado a 6 años de cárcel. Si ninguno confiesa, sólo es posible probar un delito menor (2 años de prisión a cada uno). ¿Qué hacer? La tentación por confesar es muy grande, aunque si pueden depender mutuamente, saldrían libres. Este no es un juego de suma-cero: ambos pueden perder.

Un ejemplo de juego matricial es el ***póquer***. La simulación o fingimiento (bluffing) es esencial, y para que sea efectiva debe ser impredecible. (Si su oponente descubre un patrón, usted está perdido). Las probabilidades en pro y en contra de fingir dependen de las cartas que estén a la vista, así como de las apuestas. De hecho, nuevamente el número de alternativas hace impráctico encontrar una estrategia absolutamente óptima x^*. Un buen jugador de póquer debe aproximarse bastante a x^*, y ésta puede calcularse exactamente si se acepta la siguiente simplificación enorme del juego:

X recibe una jota o un rey, con igual probabilidad, y Y siempre recibe una reina. X puede pasar y perder la apuesta de \$1, o subir la apuesta por \$2 adicionales. Si X apuesta, Y puede pasar y perder \$1, o pagar los \$2 adicionales y averiguar si X está fingiendo. Luego, la carta más alta gana los \$3 del oponente. Así, Y tiene dos posibilidades, en reacción a X (quien tiene cuatro estrategias):

Estrategias para Y
(Renglón 1) Si X apuesta, Y pasa.
(Renglón 2) Si X apuesta, Y paga los \$2 adicionales.

Estrategias para X
1) Apostar los \$2 adicionales al rey y pasar con una jota.
2) Apostar los \$2 adicionales en cualquier caso (fingir).
3) Pasar en cualquier caso, y perder \$1 (lo cual es una tontería).
4) Pasar con el rey y apostar con una jota (lo cual es una tontería).

Para calcular la matriz de pagos A se requiere un poco de paciencia:

$a_{11} = 0$: X pierde \$1 la mitad del tiempo en una jota y gana con un rey (Y pasa).
$a_{21} = 1$: En ambas apuestas X pierde \$1 la mitad del tiempo y gana \$3 la mitad del tiempo.
$a_{12} = 1$: X apuesta y Y pasa (la simulación tuvo éxito).
$a_{11} = 0$: X gana \$3 con el rey y pierde \$3 con la jota (el fingimiento fracasa).

Matriz de pagos del póquer $$A = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 1 & 0 & -1 & -2 \end{bmatrix}.$$

La estrategia óptima para X es fingir la mitad del tiempo, $x^* = \left(\frac{1}{2}, \frac{1}{2}, 0, 0\right)$. El probable perdedor Y debe escoger $y^* = \left(\frac{1}{2}, \frac{1}{2}\right)$. El valor del juego es 50 centavos para X.

Esta es una forma extraña de terminar este libro, enseñando cómo jugar un póquer suavizado (el blackjack paga mucho más). Sin embargo, me imagino que incluso el póquer tiene su sitio dentro del álgebra lineal y sus aplicaciones. Espero que haya disfrutado el libro.

Conjunto de problemas 8.5

1. ¿De qué forma son afectadas las estrategias óptimas en el juego con que empieza esta sección si los \$20 se incrementan a \$70? ¿Cuál es el valor (la ganancia media para X) de este nuevo juego?

2. Con la matriz de pagos $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, explique los cálculos por X del maximin y por Y del minimax. ¿Cuáles son las estrategias óptimas x^* y y^*?

3. Si a_{ij} es el mayor elemento en su renglón y el más pequeño en su columna, ¿X siempre escogerá la columna j y Y siempre escogerá el renglón i (sin importar el resto de la matriz)? Demuestre que el problema precedente tenía tal elemento, y luego construya A sin uno.

4. Calcule la mejor estrategia de Ys ponderando los renglones de $A = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 0 & 3 \end{bmatrix}$ con y y $1 - y$. X se centrará en la mayor de las componentes $3y + 2(1 - y)$, $4y$, y $y + 3(1 - y)$. Encuentre la más grande de estas tres (dependiendo de y), y luego encuentre la y^* entre 0 y 1 que hace que esta componente sea lo más pequeña posible.

5. Con la misma A que en el problema 4, encuentre la mejor estrategia para X. Demuestre que X sólo usa las dos columnas (la primera y la tercera) que se encuentran en el punto minimax en la gráfica.

6. Encuentre tanto las estrategias óptimas como el valor, si
$$A = \begin{bmatrix} 1 & 0 & -1 \\ -2 & -1 & 2 \end{bmatrix}$$

7. Suponga que $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. ¿Qué pesos x_1 y $1 - x_1$ proporcionan una columna de la forma $[u\ u]^T$ y qué pesos y_1 y $1 - y_1$ proporcionan un nuevo renglón $[v\ v]$? Demuestre que $u = v$.

8. Encuentre x^*, y^* y el valor v para

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

9. Calcule

$$\min_{\substack{y_i \geq 0 \\ y_1 + y_2 = 1}} \; \max_{\substack{x_i \geq 0 \\ x_1 + x_2 = 1}} (x_1 y_1 + x_2 y_2).$$

10. Explique cada una de las desigualdades en la ecuación (5). Luego, una vez que el teorema minimax la transforma en igualdades, deduzca (otra vez con palabras) las ecuaciones del punto silla (4).

11. Demuestre que $x^* = \left(\frac{1}{2}, \frac{1}{2}, 0, 0\right)$ y $y^* = \left(\frac{1}{2}, \frac{1}{2}\right)$ son estrategias óptimas en nuestra versión simplificada del póquer, calculando yAx^* y y^*Ax, y comprobando las condiciones (4) para un punto silla.

12. ¿Ya se ha demostrado que ninguna estrategia de ajedrez hace ganar a las negras? Esto es ciertamente verdadero cuando se permite que los jugadores realicen dos movimientos a la vez; si las negras tienen una estrategia ganadora, las blancas podrían adelantar y retroceder un caballo y después seguir esa estrategia, llegando a la conclusión imposible de que ambos pueden ganar.

13. Si X escoge un número primo y simultáneamente Y adivina si es par o impar (con ganancia o pérdida de \$1), ¿quién tiene la ventaja?

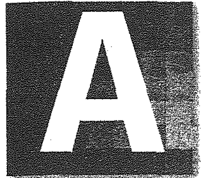
14. Si X es un mariscal de campo de fútbol americano, con la opción de correr o lanzar un pase, y Y puede defender contra una carrera o un pase, suponga que el pago (en yardas) es

$$A = \begin{bmatrix} 2 & 8 \\ 6 & -6 \end{bmatrix} \begin{array}{l} \text{defensa contra la carrera} \\ \text{defensa contra el pase.} \end{array}$$

correr pasar

¿Cuáles son las estrategias óptimas y la ganancia media en cada jugada?

Apéndice



Intersección, suma y producto de espacios

1. La intersección de dos espacios vectoriales

Al considerar dos subespacios **V** y **W** en vez de sólo uno, surgen nuevas preguntas. Primero se consideran los vectores que pertenecen a *ambos* subespacios. Esta "intersección" $\mathbf{V} \cap \mathbf{W}$ es un subespacio de estos subespacios.

> Si **V** y **W** son subespacios de un espacio vectorial, también lo es su ***intersección*** **$\mathbf{V} \cap \mathbf{W}$. Los vectores que pertenecen tanto a V como a W constituyen un subespacio.**

Suponga que x y y son vectores que están en **V** y también en **W**. Debido a que **V** y **W** son espacios vectoriales por derecho propio, $x + y$ y cx están en **V** y en **W**. ***Los resultados de la suma y la multiplicación escalar también están en la intersección.***

Dos planos que pasan por el origen (o dos "hiperplanos" en $\mathbf{R}^n$) se cortan en un subespacio. La intersección de varios subespacios, o de un número infinito también es un subespacio.

Ejemplo 1 La intersección de dos subespacios *ortogonales* **V** y **W** es el subespacio $\mathbf{V} \cap \mathbf{W} = \{0\}$ que consta de un solo punto. Sólo el vector cero es ortogonal a sí mismo.

Ejemplo 2 Suponga que **V** y **W** son los espacios de matrices triangulares superior e inferior de n por n. La intersección $\mathbf{V} \cap \mathbf{W}$ es el conjunto de *matrices diagonales*, que pertenecen a ambos subespacios triangulares. El resultado de sumar matrices diagonales, o de multiplicarlas por c, es una matriz diagonal.

Ejemplo 3 Suponga que **V** es el espacio nulo de A y que **W** es el espacio nulo de B. Entonces $\mathbf{V} \cap \mathbf{W}$ es el menor espacio nulo de la matriz más grande C:

Intersección de espacios nulos $\quad N(A) \cap N(B)$ **es el espacio nulo de** $C = \begin{bmatrix} A \\ B \end{bmatrix}$.

$Cx = 0$ requiere tanto $Ax = 0$ como $Bx = 0$. Así, x tiene que pertenecer a ambos espacios nulos.

2. La suma de dos espacios vectoriales

Casi siempre, después de analizar la intersección de dos conjuntos, resulta natural estudiar su unión. Con espacios vectoriales esto no es natural. *La unión* $\mathbf{V} \cup \mathbf{W}$ *de dos subespacios en general no es un subespacio.* Si **V** y **W** son el eje x y el eje y en el plano, los dos ejes juntos no son un subespacio. La suma de (1, 0) y (0, 1) no está en ninguno de estos ejes.

Se desea combinar **V** y **W**. En vez de estudiar su unión, lo hacemos con su suma.

DEFINICIÓN Si **V** y **W** son subespacios de un espacio dado, también lo es su ***suma***. $\mathbf{V} + \mathbf{W}$ contiene a todas las combinaciones $v + w$, donde v está en **V** y w está en **W**.

$\mathbf{V} + \mathbf{W}$ es el menor espacio vectorial que contiene tanto a **V** como a **W**. La **suma** del eje x y del eje y es todo el plano x-y, como también lo es la suma de dos rectas distintas cualesquiera, *sean perpendiculares o no.* Si **V** es el eje x y **W** es la recta $x = y$ a 45°, entonces cualquier vector como (5, 3) puede separarse en $v + w = (2, 0) + (3, 3)$. Así, $\mathbf{V} + \mathbf{W}$ es todo $\mathbf{R}^2$.

Ejemplo 4 Suponga que **V** y **W** son complementos ortogonales en $\mathbf{R}^n$. Entonces su suma es $\mathbf{V} + \mathbf{W} = \mathbf{R}^n$. Todo x es la suma de sus proyecciones en **V** y **W**.

Ejemplo 5 Si **V** es el espacio de matrices triangulares superiores y **W** es el espacio de matrices triangulares inferiores, entonces $\mathbf{V} + \mathbf{W}$ es el espacio de *todas* las matrices. Toda matriz de n por n puede escribirse como la suma de una matriz triangular superior y una matriz triangular inferior —de muchas formas, ya que las diagonales no están determinadas de manera única.

Estos subespacios triangulares tienen dimensión $n(n + 1)/2$. El espacio $\mathbf{V} + \mathbf{W}$ de todas las matrices tiene dimensión n^2. El espacio $\mathbf{V} \cap \mathbf{W}$ de las matrices diagonales tiene dimensión n. La fórmula (8) que se presenta a continuación se convierte en $n^2 + n = n(n + 1)/2 + n(n + 1)/2$.

Ejemplo 6 Si **V** es el espacio columna de A y **W** es el espacio columna de B, entonces $\mathbf{V} + \mathbf{W}$ ***es el espacio columna de la matriz más grande*** $[A \;\; B]$. La dimensión de $\mathbf{V} + \mathbf{W}$ puede ser menor que las dimensiones combinadas de **V** y **W** (porque estos dos espacios podrían traslaparse):

Suma de espacios columna $\quad \dim(\mathbf{V} + \mathbf{W}) = \text{rango de } [A \;\; B]. \quad (6)$

El *cálculo de* $\mathbf{V} \cap \mathbf{W}$ *es más sutil.* Para la intersección de espacios columna, un buen método consiste en poner bases de **V** y **W** en las columnas de A y B. El espacio nulo de $[A \;\; B]$ conduce a $\mathbf{V} \cap \mathbf{W}$ (consulte el problema 9). *Estos espacios tienen la misma dimensión* (la nulidad de $[A \; B]$). Al combinar con $\dim(\mathbf{V} + \mathbf{W})$ se obtiene

$$\dim(\mathbf{V} + \mathbf{W}) + \dim(\mathbf{V} \cap \mathbf{W}) = \text{rango de } [A \;\; B] + \text{nulidad de } [A \;\; B]. \quad (7)$$

Se sabe que el rango más la nulidad (contando las columnas pivote más las columnas libres) siempre es igual al número total de columnas. Cuando $[A \; B]$ tiene $k + \ell$ columnas, con $k = \dim \mathbf{V}$ y $\ell = \dim \mathbf{W}$, se llega a una conclusión clara:

Dimensión fórmula $\quad \dim(\mathbf{V} + \mathbf{W}) + \dim(\mathbf{V} \cap \mathbf{W}) = \dim(\mathbf{V}) + \dim(\mathbf{W}). \quad (8)$

Que no es una mala fórmula. El traslape de **V** y **W** es $\mathbf{V} \cap \mathbf{W}$.

3. Producto cartesiano de dos espacios vectoriales

Si **V** tiene dimensión n y **W** tiene dimensión q, su producto cartesiano $\mathbf{V} \times \mathbf{W}$ tiene dimensión $n + q$.

DEFINICIÓN $\mathbf{V} \times \mathbf{W}$ ***contiene a todas las parejas de vectores*** $x = (v, w)$.

Al sumar (v, w) a (v^*, w^*) en este espacio producto se obtiene $(v + v^*, w + w^*)$, y al multiplicar por c se obtiene (cv, cw). Todas las operaciones en $\mathbf{V} \times \mathbf{W}$ se realizan componente por componente.

Ejemplo 7 El producto cartesiano de $\mathbf{R}^2$ y $\mathbf{R}^3$ se parece bastante a $\mathbf{R}^5$. Un vector típico x en $\mathbf{R}^2 \times \mathbf{R}^3$ es ((1, 2), (4, 6, 5)); un vector de $\mathbf{R}^2$ y uno de $\mathbf{R}^3$. Esto se ve como (1, 2, 4, 6, 5) en $\mathbf{R}^5$.

Los productos cartesianos están asociados de forma natural con las *matrices en bloque*. De $\mathbf{R}^5$ a $\mathbf{R}^5$ se tienen matrices ordinarias de 5 por 5. En el espacio producto $\mathbf{R}^2 \times \mathbf{R}^3$, la forma natural de una matriz en bloque M de 5 por 5 es:

$$M = \begin{bmatrix} \mathbf{R}^2 \text{ a } \mathbf{R}^2 & \mathbf{R}^3 \text{ a } \mathbf{R}^2 \\ \mathbf{R}^2 \text{ a } \mathbf{R}^3 & \mathbf{R}^3 \text{ a } \mathbf{R}^3 \end{bmatrix} = \begin{bmatrix} 2 \text{ por } 2 & 2 \text{ por } 3 \\ 3 \text{ por } 2 & 3 \text{ por } 3 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

El resultado de la multiplicación de una matriz por un vector es $(Av + Bw, Cv + Dw)$, lo cual no es muy fascinante.

4. Producto tensorial de dos espacios vectoriales

De alguna manera se quiere contar con un espacio producto cuya dimensión sea n *multiplicada por* q. **Los vectores en este "espacio tensorial" (que se denota por ⊗), se parecen a las matrices de** n **por** q. Para el producto tensorial $\mathbf{R}^2 \otimes \mathbf{R}^3$, los vectores se parecen a las matrices de 2 por 3. La dimensión de $\mathbf{R}^2 \times \mathbf{R}^3$ es 5, pero la dimensión de $\mathbf{R}^2 \otimes \mathbf{R}^3$ es 6.

Se empieza con $v = (1, 2)$ y $w = (4, 6, 5)$ en $\mathbf{R}^2$ y $\mathbf{R}^3$. El producto cartesiano simplemente los aproxima mutuamente como (v, w). El producto tensorial combina a v y w en **la matriz** vw^{T} ***de rango*** **1:**

Columna multiplicada por renglón $\quad v \otimes w = vw^{\mathrm{T}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 4 & 6 & 5 \end{bmatrix} = \begin{bmatrix} 4 & 6 & 5 \\ 8 & 12 & 10 \end{bmatrix}.$

Todas las matrices especiales vw^{T} pertenecen al producto tensorial $\mathbf{R}^2 \otimes \mathbf{R}^3$. El espacio producto es *generado* por estos vectores $v \otimes w$. Las combinaciones de las matrices de rango 1 proporcionan *todas* las matrices de 2 por 3, por lo que la dimensión de $\mathbf{R}^2 \otimes \mathbf{R}^3$ es 6. En términos abstractos: el producto tensorial $\mathbf{V} \otimes \mathbf{W}$ se identifica con el espacio de transformaciones lineales que van de **V** a **W**.

Si **V** es sólo una recta en $\mathbf{R}^2$ y **V** es sólo una recta en $\mathbf{R}^3$, entonces $\mathbf{V} \otimes \mathbf{W}$ es sólo una "recta en el espacio de matrices". Ahora las dimensiones son $1 \times 1 = 1$. Todas las matrices vw^{T} de rango 1 son múltiplos de una matriz.

Bases del producto tensorial. Cuando **V** es $\mathbf{R}^2$ y **W** es $\mathbf{R}^3$, se tiene una base estándar para todas las matrices de 2 por 3 (un espacio de seis dimensiones):

Base $\quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$

La base para $\mathbf{R}^2 \otimes \mathbf{R}^3$ se construyó de forma natural. Se comenzó con la base estándar $v_1 = (1, 0)$ y $v_2 = (0, 1)$ para $\mathbf{R}^2$. Estos vectores se combinaron con los vectores básicos $w_1 = (1, 0, 0)$, $w_2 = (0, 1, 0)$ y $w_3 = (0, 0, 1)$ en $\mathbf{R}^3$. Cada pareja $v_i \otimes w_j$ corresponde a uno de los seis vectores básicos (las matrices de 2 por 3 de antes) en el producto tensorial $\mathbf{V} \otimes \mathbf{W}$. Esta construcción también tiene éxito para subespacios:

Base: Suponga que $\mathbf{V}$ y $\mathbf{W}$ son subespacios de $\mathbf{R}^m$ y $\mathbf{R}^p$ con bases $v_1, \ldots, v_n$ y $w_1, \ldots, w_q$. Entonces las matrices nq de rango -1 $v_i w_j^T$ constituyen una base para $\mathbf{V} \otimes \mathbf{W}$.

$\mathbf{V} \otimes \mathbf{W}$ es un subespacio de dimensión nq de las matrices de m por p. Un algebrista haría corresponder esta construcción matricial con la definición abstracta de $\mathbf{V} \otimes \mathbf{W}$. Luego, los productos tensoriales pueden ir más allá del caso específico de los vectores columna.

5. Producto de Kronecker $A \otimes B$ de dos matrices

Una matriz A de m por n transforma cualquier vector v de $\mathbf{R}^n$ en un vector Av de $\mathbf{R}^m$. De manera semejante, una matriz B de p por q transforma w en Bw. Las dos matrices juntas transforman vw^T en Avw^TB^T. Esta es una transformación lineal (de productos tensoriales) y debe provenir de una matriz.

¿Cuál es el tamaño de la matriz $A \otimes B$? Ésta lleva el espacio nq-dimensional $\mathbf{R}^n \otimes \mathbf{R}^q$ en el espacio mp-dimensional $\mathbf{R}^m \otimes \mathbf{R}^p$. En consecuencia, la forma de la matriz es de mp por nq. Este producto de Kronecker que también se denomina producto tensorial) se escribe

Producto de Kronecker mp renglones, nq columnas

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}. \tag{9}$$

¡Observe la estructura especial de esta matriz! Muchas matrices en bloque importantes tienen esta forma de Kronecker. Aquéllas provienen de aplicaciones bidimensionales, donde A es una "matriz en la dirección x" y B actúa en la dirección y (a continuación se proporcionan algunos ejemplos). Si A y B son cuadradas, también lo son $m = n$ y $p = q$, entonces la matriz grande $A \otimes B$ también es cuadrada.

Ejemplo 8 **(Diferencias finitas en las direcciones x y y)** La ecuación diferencial original de Laplace $-\partial^2 u/\partial x^2 - \partial^2 u/\partial y^2 = 0$ se sustituye por diferencias finitas a fin de encontrar valores para u en una malla bidimensional. Las diferencias en la dirección x se suman a las diferencias en la dirección y, uniendo cinco valores vecinos de u:

−1 2 −1 + −1 2 −1 ⟶ −1 4 −1 (−1 arriba y abajo)

$$-u_{i+1,j} + 2u_{i,j} - u_{i-1,j}$$
$$-u_{i,j+1} + 2u_{i,j} - u_{i,j-1}$$
$$= 0$$

diferencias-x diferencias-y suma

Una ecuación de 5 puntos está centrada en cada uno de los nueve puntos de la malla. La matriz de 9 por 9 (que se denomina A_{2D}) se construye a partir de la matriz "1D" de 3 por 3 para las diferencias a lo largo de una recta:

Matriz de diferencias en una dirección $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ **Matriz identidad en otra dirección** $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Los productos de Kronecker producen tres diferencias 1D a lo largo de tres rectas, hacia arriba o transversales:

Una dirección $$A \otimes I = \begin{bmatrix} 2I & -I & 0 \\ -I & 2I & -I \\ 0 & -I & 2I \end{bmatrix}.$$

Otra dirección $$I \otimes A = \begin{bmatrix} A & 0 & 0 \\ 0 & A & 0 \\ 0 & 0 & A \end{bmatrix}.$$

Ambas direcciones $$A_{2D} = (A \otimes I) + (I \otimes A) = \begin{bmatrix} A + 2I & -I & 0 \\ -I & A + 2I & -I \\ 0 & -I & A + 2I \end{bmatrix}.$$

La suma $(A \otimes I) + (I \otimes A)$ es la matriz de 9 por 9 de Laplace para la ecuación en diferencias de 5 puntos (en la sección 1.7 era para 1D y en la sección 7.4 se mencionó a 2D). El renglón de en medio de esta matriz de 9 por 9 muestra todos los cinco elementos diferentes de cero para la molécula de 5 puntos:

Lejos de la frontera Renglón 5 de $A_{2D} = [0 \ \ -1 \ \ 0 \quad -1 \ \ 4 \ \ -1 \quad 0 \ \ -1 \ \ 0]$.

Ejemplo 9 **(Matriz de Fourier en 2D)** La matriz unidimensional F de Fourier es la matriz compleja más importante en el mundo. La Transformada Rápida de Fourier en la sección 3.5 es una forma rápida para multiplicar por esa matriz F. Así, la TRF transforma el "dominio temporal en el dominio de frecuencias" para una señal de audio en 1D. **Para imágenes, se requiere la transformada 2D:**

Matriz de Fourier en 2D $F_{2D} = F \otimes F =$ **Se transforma a lo largo de cada renglón, y luego hacia abajo de cada columna**

La imagen es un arreglo bidimensional de valores de pixeles. Es transformada por F_{2D} en un arreglo bidimensional de coeficientes de Fourier. Este arreglo puede comprimirse, transmitirse y almacenarse. Luego, *la transformada inversa* regresa de los coeficientes de Fourier a los valores de pixeles. Es necesario conocer la regla inversa para los productos de Kronecker:

La inversa de la matriz $A \otimes B$ es la matriz $A^{-1} \otimes B^{-1}$.

¡La TRF también acelera la transformada inversa 2D! Simplemente se invierte en una dirección seguida de la otra dirección. Se está sumando $\sum\sum c_{k\ell} e^{ikx} e^{i\ell y}$ sobre k y luego sobre ℓ.

La matriz de diferencias de Laplace $A_{2D} = (A \otimes I) + (I \otimes A)$ no tiene una fórmula inversa sencilla. Es por ello que la ecuación $A_{2D}u = b$ ha sido estudiada con tanto cuidado. Uno de los métodos más rápidos es diagonalizar A_{2D} usando la matriz de vectores característicos (que es la matriz seno $S \otimes S$ de Fourier, muy semejante a A_{2D}). Los valores característicos de A_{2D} provienen inmediatamente de los valores característicos de A_{1D}:

Los n^2 valores característicos de $(A \otimes I) + (I \otimes B)$ son todas las sumas $\lambda_i(A) + \lambda_j(B)$. Los n^2 valores característicos de $A \otimes B$ son todos los productos $\lambda_i(A)\lambda_j(B)$.

Si A y B son de n por n, el determinante de $A \otimes B$ (el producto de sus valores característicos) es $(\det A)^n(\det B)^n$. La traza de $A \otimes B$ es (traza A)(traza B). ¡En este apéndice se ilustran tanto el "álgebra lineal pura" como sus aplicaciones cruciales!

Conjunto de problemas A

1. Suponga que **S** y **T** son subespacios de $\mathbf{R}^{13}$, con dim **S** = 7 y dim **T** = 8.

a) ¿Cuál es la máxima dimensión posible de $\mathbf{S} \cap \mathbf{T}$?
b) ¿Cuál es la mínima dimensión posible de $\mathbf{S} \cap \mathbf{T}$?
c) ¿Cuál es la mínima dimensión posible de $\mathbf{S} + \mathbf{T}$?
d) ¿Cuál es la máxima dimensión posible de $\mathbf{S} + \mathbf{T}$?

2. ¿Cuáles son las intersecciones de las siguientes parejas de subespacios?

a) El plano x-y y el plano z-y en $\mathbf{R}^3$.
b) La recta que pasa por (1, 1, 1) y el plano que pasa por (1, 0, 0) y (0, 1, 1).
c) El vector cero y todo el espacio $\mathbf{R}^3$.
d) El plano S perpendicular a (1, 1, 0) y (0, 1, 1) en $\mathbf{R}^3$.

¿Cuáles son las *sumas* de estas parejas de subespacios?

3. En el espacio de todas las matrices de 4 por 4, sea **V** el subespacio de las matrices *tridiagonales* y **W** el subespacio de las matrices *triangulares superiores*. Describa el subespacio **V** + **W**, cuyos elementos son las matrices superiores de Hessenberg. ¿Cuál es $\mathbf{V} \cap \mathbf{W}$? Compruebe la fórmula (8).

4. Si $\mathbf{V} \cap \mathbf{W}$ contiene sólo al vector cero, entonces la ecuación (8) se convierte en dim(**V** + **W**) = dim **V** + dim **W**. Compruebe este hecho cuando **V** es el espacio renglón de A, **W** es el espacio nulo de A y la matriz A es de m por n de rango r. ¿Cuáles son las dimensiones?

5. Proporcione un ejemplo en $\mathbf{R}^3$ para el que $\mathbf{V} \cap \mathbf{W}$ contiene sólo al vector cero, pero **V** no es ortogonal a **W**.

6. Si $\mathbf{V} \cap \mathbf{W} = \{0\}$, entonces **V** + **W** se denomina *suma directa* de **V** y **W**, con la notación especial $\mathbf{V} \otimes \mathbf{W}$. Si **V** es generado por (1, 1, 1) y (1, 0, 1), escoja un subespacio **W** de modo que $\mathbf{V} \otimes \mathbf{W} = \mathbf{R}^3$. Explique por qué cualquier vector x en la suma directa $\mathbf{V} \otimes \mathbf{W}$ puede escribirse de una *y sólo una* forma como $x = v + w$ (con v en **V** y w en **W**).

7. Encuentre una base para la suma **V** + **W** del espacio **V** generado por $v_1 = (1, 1, 0, 0)$, $v_2 = (1, 0, 1, 0)$ y el espacio **W** generado por $w_1 = (0, 1, 0, 1)$, $w_2 = (0, 0, 1, 1)$. También encuentre la dimensión de $\mathbf{V} \cap \mathbf{W}$ y una base para éste.

8. A partir de la ecuación (8), demuestre que rango $(A + B) \leq$ rango(A) + rango(B).

9. **La intersección de $C(A) \cap C(B)$ coincide con el espacio nulo de** $[A \quad B]$. Cada $y = Ax_1 = Bx_2$ en los espacios columna tanto de A como de B coinciden con $x = (x_1, -x_2)$ en el espacio nulo, porque $[A \quad B]x = Ax_1 - Bx_2 = 0$. Compruebe que $y = (6, 3, 6)$ coincide con $x = (1, 1, -2, -3)$ y encuentre la intersección $C(A) \cap C(B)$, para

$$A = \begin{bmatrix} 1 & 5 \\ 3 & 0 \\ 2 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & 0 \\ 0 & 1 \\ 0 & 2 \end{bmatrix}.$$

10. Multiplique $A \otimes B$ por $A^{-1} \otimes B^{-1}$ para obtener $AA^{-1} \otimes BB^{-1} = I \otimes I = I_{2D}$.

11. ¿Cuál es la matriz de Fourier de 4 por 4 $F_{2D} = F \otimes F$ para $F = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$?

12. Suponga que $Ax = \lambda(A)x$ y $By = \lambda(B)y$. Forme un vector columna largo z con n^2 componentes, x_1y, luego x_2y y por último x_ny. Demuestre que z es un vector característico de $(A \otimes I)z = \lambda(A)z$ y $(A \otimes B)z = \lambda(A)\lambda(B)z$.

13. ¿Cuál debe ser la matriz de Laplace de 7 puntos para $-u_{xx} - u_{yy} - u_{zz} = 0$? Esta matriz "tridimensional" se construye a partir de productos de Kronecker usando I y A_{1D}.

Apéndice

B La forma de Jordan

Dada una matriz cuadrada A, se quiere escoger M de forma que $M^{-1}AM$ sea lo más diagonal posible. En el caso más sencillo, A tiene un conjunto completo de vectores característicos que se convierten en las columnas de M, conocida como S. La forma de Jordan es $J = M^{-1}AM = \Lambda$; se construyó completamente a partir de bloques $J_i = \lambda_i$ de 1 por 1, y el objeto de una matriz diagonal se ha alcanzado por completo. En el caso más general y difícil, faltan algunos vectores característicos y una forma diagonal es imposible. Ese caso constituye ahora nuestro principal interés.

Se repite el teorema que debe demostrarse:

> Si una matriz A tiene s vectores característicos linealmente independientes, entonces es semejante a una matriz J que es la ***forma de Jordan***, con s bloques cuadrados en la diagonal:
>
> $$J = M^{-1}AM = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{bmatrix}.$$
>
> Cada bloque tiene un vector característico, un valor característico y unos justo arriba de la diagonal:
>
> $$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \cdot & \cdot & \\ & & \cdot & 1 \\ & & & \lambda_i \end{bmatrix}.$$

Un ejemplo de esta forma de Jordan es el siguiente:

$$J = \begin{bmatrix} 8 & 1 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 8 & 1 \\ 0 & 8 \end{bmatrix} & & \\ & \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} & \\ & & [0] \end{bmatrix} = \begin{bmatrix} J_1 & & \\ & J_2 & \\ & & J_3 \end{bmatrix}.$$

El valor característico doble $\lambda = 8$ sólo tiene un simple vector característico, en la primera dirección de coordenadas $e_1 = (1, 0, 0, 0, 0)$; como resultado, $\lambda = 8$ sólo aparece en un simple bloque J_1. El valor característico triple $\lambda = 0$ tiene dos vectores característicos, e_3 y e_5, que corresponden a los dos bloques de Jordan J_2 y J_3. Si A tuviese 5 vectores característicos, todos los bloques serían de 1 por 1 y J sería diagonal.

La pregunta clave es: *Si A es alguna otra matriz de 5 por 5, ¿en qué condiciones su forma de Jordan sería esta misma J? ¿Cuándo existe una M tal que $M^{-1}AM = J$?* Como primer requerimiento, cualquier matriz semejante A debe compartir los mismos valores característicos 8, 8, 0, 0, 0. Sin embargo, la matriz diagonal con estos valores característicos

no es semejante a J, y la pregunta planteada en realidad concierne a los vectores característicos.

Para contestar, $M^{-1}AM = J$ se escribe en forma más simple como $AM = MJ$:

$$A \begin{bmatrix} & & & & \\ x_1 & x_2 & x_3 & x_4 & x_5 \\ & & & & \end{bmatrix} = \begin{bmatrix} & & & & \\ x_1 & x_2 & x_3 & x_4 & x_5 \\ & & & & \end{bmatrix} \begin{bmatrix} 8 & 1 & & & \\ 0 & 8 & & & \\ & & 0 & 1 & \\ & & 0 & 0 & \\ & & & & 0 \end{bmatrix}.$$

Al efectuar la multiplicación columna por columna,

$$Ax_1 = 8x_1 \quad \text{y} \quad Ax_2 = 8x_2 + x_1 \tag{10}$$

$$Ax_3 = 0x_3 \quad \text{y} \quad Ax_4 = 0x_4 + x_3 \quad \text{y} \quad Ax_5 = 0x_5. \tag{11}$$

Ahora es posible reconocer las condiciones sobre A. Debe tener tres vectores característicos genuinos, así como J. El que cumple $\lambda = 8$ debe ir en la primera columna de M, exactamente como si hubiese ido en la primera columna de S: $Ax_1 = 8x_1$. Los otros dos, que se identifican como x_3 y x_5, van en las columnas tercera y quinta de M: $Ax_3 = Ax_5 = 0$. Por último, debe haber otros dos vectores especiales: los *vectores característicos generalizados* x_2 y x_4. Se considera que x_2 pertenece a una ***cadena de vectores***, encabezados por x_1 y descrita por las ecuaciones (10). De hecho, x_2 es el único otro vector en la cadena, y el bloque J_1 correspondiente es de orden 2. La ecuación (11) describe *dos cadenas distintas*, una en la que x_4 sigue a x_3, y otra en la que x_5 está solo: los bloques J_2 y J_3 son de 2 por 2 y de 1 por 1.

La búsqueda de la forma de Jordan de A se convierte en la búsqueda de estas cadenas de vectores, cada una encabezada por un vector característico: Para toda i,

$$\text{ya sea} \quad Ax_i = \lambda_i x_i \quad \text{o bien,} \quad Ax_i = \lambda_i x_i + x_{i-1}. \tag{12}$$

Los vectores x_i van en las columnas de M, y cada cadena produce un simple bloque en J. Esencialmente, es necesario demostrar que estas cadenas pueden construirse para cada matriz A. Entonces, si las cadenas coinciden con las ecuaciones particulares (10) y (11), la J obtenida será la forma de Jordan de A.

Considero que la idea de Filippov hace de la construcción lo más clara y sencilla posible.* Procede por inducción matemática, empezando con el hecho de que toda matriz de 1 por 1 ya está en forma de Jordan. Puede suponerse que la construcción se logra para todas las matrices de orden menor que n —esta es la "hipótesis de inducción"— y luego explicar los pasos para una matriz de orden n. Hay tres pasos, que después de una descripción general se aplican a un ejemplo específico.

Paso 1 Si se supone que A es singular, entonces su espacio columna tiene dimensión $r < n$. Observando sólo dentro de este espacio más pequeño, la hipótesis de inducción garantiza que una forma de Jordan es posible: en el espacio columna debe haber r vectores independientes w_i tales que

$$\text{ya sea} \quad Aw_i = \lambda_i w_i \quad \text{o bien,} \quad Aw_i = \lambda_i w_i + w_{i-1}. \tag{13}$$

Paso 2 Suponga que el espacio nulo y el espacio columna de A tienen una intersección de dimensión p. Por supuesto, todo vector en el espacio nulo es un vector característico correspondiente a $\lambda = 0$. Por consiguiente, en el paso 1 debe haber p cadenas que comenzaron en este valor característico, y se tiene interés en los vectores w_i que

* A. F. Filippov, A short proof of the reduction to Jordan form, Moscow Univ. Math. Bull., volume 26(1971) pp. 70-71.

vienen al final de estas cadenas. Cada uno de estos p vectores está en el espacio columna, por lo que cada uno es una combinación de las columnas de A: $w_i = Ay_i$ para alguna y_i.

Paso 3 La dimensión del espacio nulo siempre es $n - r$. En consecuencia, de forma independiente con respecto a su intersección p-dimensional con el espacio columna, debe contener $n - r - p$ vectores básicos adicionales z_i *fuera* de esa intersección.

A continuación se reúnen estos pasos para obtener el teorema de Jordan:

Los r vectores w_i, los p vectores y_i y los $n - r - p$ vectores z_i forman cadenas de Jordan para la matriz A, y estos vectores son linealmente independientes. Van en las columnas de M, y $J = M^{-1}AM$ está en forma de Jordan.

Si se desea renumerar estos vectores como $x_1, \ldots, x_n$ y hacerlos coincidir con la ecuación (12), entonces cada y_i debe insertarse de inmediato después del w_i del que provino; completa una cadena en la que $\lambda_i = 0$. Las zs vienen hasta el final, cada una sola en su propia cadena; de nuevo el valor característico es cero, ya que las zs están en el espacio nulo. Los bloques con valores característicos diferentes de cero ya se habían terminado en el paso 1, los bloques con valores característicos cero aumentan por un renglón y una columna en el paso 2 y el paso 3 contribuye con cualesquiera bloques de 1 por 1 $J_i = [0]$.

A continuación se intentará trabajar un ejemplo, y a fin de permanecer próximo a las páginas previas, los valores característicos se toman como 8, 8, 0, 0, 0:

$$A = \begin{bmatrix} 8 & 0 & 0 & 8 & 8 \\ 0 & 0 & 0 & 8 & 8 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 \end{bmatrix}.$$

Paso 1 La dimensión del espacio columna es $r = 3$, y es generado por los vectores de coordenadas e_1, e_2, e_5. Para buscar dentro del espacio se ignoran los renglones tercero y cuarto y las columnas de A, lo que queda tiene valores característicos 8, 8, 0 y su forma de Jordan proviene de los vectores

$$w_1 = \begin{bmatrix} 8 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad w_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad w_3 = \begin{bmatrix} 0 \\ 8 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Los w_i están en el espacio columna, completan la cadena para $\lambda = 8$ e inician la cadena para $\lambda = 0$:

$$Aw_1 = 8w_1, \qquad Aw_2 = 8w_2 + w_1, \qquad Aw_3 = 0w_3. \tag{14}$$

Paso 2 El espacio nulo de A contiene a e_2 y e_3, por lo que su intersección con el espacio columna es generada por e_2. Por consiguiente, $p = 1$ y, como era de esperar, en la ecuación (14) hay una cadena correspondiente a $\lambda = 0$. El vector w_3 viene al final (y también al principio) de esa cadena, y $w_3 = A(e_4 - e_1)$. En consecuencia, $y = e_4 - e_1$.

Paso 3 El ejemplo tiene $n - r - p = 5 - 3 - 1 = 1$, y $z = e_3$ está en el espacio nulo pero fuera del espacio columna. Es esta z la que produce un bloque de 1 por 1 en J.

Si los cinco vectores se ensamblan, las cadenas completas son

$$Aw_1 = 8w_1, \quad Aw_2 = 8w_2 + w_1, \quad Aw_3 = 0w_3, \quad Ay = 0y + w_3, \quad Az = 0z.$$

Al comparar con las ecuaciones (10) y (11), se tiene una correspondencia perfecta: la forma de Jordan del ejemplo es exactamente la J que se escribió previamente. Al escribir los cinco vectores en las columnas de M debe obtenerse $AM = MJ$, o bien, $M^{-1}AM = J$:

$$M = \begin{bmatrix} 8 & 0 & 0 & -1 & 0 \\ 0 & 1 & 8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

Confiamos bastante en las matemáticas (o somos demasiado flojos), por lo que no multiplicamos $M^{-1}AM$.

En la construcción de Filippov, la única cuestión técnica es comprobar la independencia de toda la colección w_i, y_i y z_i. En consecuencia, se supone que alguna combinación es cero:

$$\sum c_i w_i + \sum d_i y_i + \sum g_i z_i = 0. \tag{15}$$

Al multiplicar por A, y usar las ecuaciones (13) para las w_i, así como $Az_i = 0$,

$$\sum c_i \begin{bmatrix} \lambda_i w_i \\ \text{o bien,} \\ \lambda_i w_i + w_{i-1} \end{bmatrix} + \sum d_i A y_i = 0. \tag{16}$$

Las Ay_i son las w_i especiales al final de las cadenas correspondientes a $\lambda_i = 0$, de modo que no pueden aparecer en la primera suma. (Están multiplicadas por cero en $\lambda_i w_i$.) Debido a que la ecuación (16) es alguna combinación de las w_i, que por la hipótesis de inducción eran independientes —proporcionaban la forma de Jordan en el espacio columna— se concluye que *cada d_i debe ser cero*. Volviendo a la ecuación (15), esto deja $\sum c_i w_i = -\sum g_i z_i$ y el miembro izquierdo está en el espacio columna. Debido a que las zs eran independientes de ese espacio, cada g_i debe ser cero. Por último, $\sum c_i w_i = 0$ y la independencia de las w_i produce $c_i = 0$.

Si la A original no hubiese sido singular, entonces los tres pasos se hubieran aplicado a $A' = A - cI$. (La constante c se escoge para hacer singular a A', y puede ser cualquiera de los valores característicos de A.) El algoritmo pone A' en su forma de Jordan $M^{-1}A'M = J'$ al producir las cadenas x_i a partir de las w_i, y_i y z_i. Luego, la forma de Jordan para A utiliza las mismas cadenas y la misma M:

$$M^{-1}AM = M^{-1}A'M + M^{-1}cM = J' + cI = J.$$

Con esto se completa la demostración de que toda A es semejante a alguna matriz de Jordan J. Salvo un reagrupamiento de los bloques, ***es semejante sólo a una J así***; hay una forma de Jordan única para A. Así, el conjunto de todas las matrices se separa en una cantidad de familias con la siguiente propiedad: *todas las matrices de la misma familia tienen la misma forma de Jordan, y todas son semejantes entre sí* (y a J), *pero ningunas matrices pertenecientes a familias distintas son semejantes*. En cada familia, J es la más hermosa, en caso de que a usted le agrade que las matrices sean casi diagonales. Con esta clasificación de familias terminamos.

Ejemplo 1

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{con } \lambda = 0, 0, 0.$$

Esta matriz tiene rango $r = 2$ y sólo un vector característico. Dentro del espacio columna hay una sola cadena w_1, w_2 que coincide con las dos últimas columnas:

$$A\begin{bmatrix}1\\0\\0\end{bmatrix} = 0 \quad \text{y} \quad A\begin{bmatrix}2\\1\\0\end{bmatrix} = \begin{bmatrix}1\\0\\0\end{bmatrix},$$

o bien,

$$Aw_1 = 0 \quad \text{y} \quad Aw_2 = 0w_2 + w_1.$$

El espacio nulo está completamente dentro del espacio columna, y es generado por w_1. En consecuencia, $p = 1$ en el paso 2, y el vector y proviene de la ecuación

$$Ay = w_2 = \begin{bmatrix}2\\1\\0\end{bmatrix}, \quad \text{cuya solución es} \quad y = \begin{bmatrix}0\\0\\1\end{bmatrix}.$$

Finalmente, la cadena w_1, w_2, y va en la matriz M:

$$M = \begin{bmatrix}1&2&0\\0&1&0\\0&0&1\end{bmatrix}, \quad \text{y} \quad M^{-1}AM = \begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix} = J.$$

Aplicación a $du/dt = Au$

Como siempre, el problema se simplifica separando las incógnitas. Esta separación está completa sólo cuando hay un conjunto completo de valores característicos y $u = Sv$; el mejor cambio de variables en este caso es $u = Mv$. Esto produce la nueva ecuación $M\,dv/dt = AMv$, o $dv/dt = Jv$, que es lo más simple como permiten las circunstancias. Está acoplada sólo por los 1s fuera de la diagonal dentro de cada bloque de Jordan. En el ejemplo precedente, que tenía un solo bloque, $du/dt = Au$ se convierte en

$$\frac{dv}{dt} = \begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}v \quad \text{o bien,} \quad \begin{matrix}da/dt = b\\ db/dt = c\\ dc/dt = 0\end{matrix} \quad \text{o bien,} \quad \begin{matrix}a = a_0 + b_0t + c_0t^2/2\\ b = \quad b_0 + c_0t\\ c = \quad c_0.\end{matrix}$$

El sistema se resuelve trabajando hacia arriba a partir de la última ecuación, y en cada paso entra una nueva potencia de t. (Un bloque de ℓ por ℓ tiene potencias tan altas como $t^{\ell-1}$). Las exponenciales de J, en este caso y en el ejemplo anterior de 5 por 5, son

$$e^{Jt} = \begin{bmatrix}1&t&t^2/2\\0&1&t\\0&0&1\end{bmatrix} \quad \text{y} \quad \begin{bmatrix}e^{8t}&te^{8t}&0&0&0\\0&e^{8t}&0&0&0\\0&0&1&t&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix}.$$

Puede verse cómo los coeficientes de a, b y c aparecen en la primera exponencial. Y en el segundo ejemplo es posible identificar a todas las cinco de las "soluciones especiales" de $du/dt = Au$. Tres de ellas son las exponenciales puras $u_1 = e^{8t}x_1$, $u_3 = e^{0t}x_3$ y $u_5 = e^{0t}x_5$, formadas como de costumbre a partir de los tres vectores característicos de A. Las otras dos implican los vectores característicos generalizados x_2 y x_4:

$$u_2 = e^{8t}(tx_1 + x_2) \quad \text{y} \quad u_4 = e^{0t}(tx_3 + x_4). \tag{17}$$

La solución más general de $du/dt = Au$ es una combinación de $c_1u_1 + \cdots + c_5u_5$, y la combinación que coincide con u_0 en el instante $t = 0$ de nuevo es

$$u_0 = c_1x_1 + \cdots + c_5x_5, \quad \text{o bien} \quad u_0 = Mc, \quad \text{o bien} \quad c = M^{-1}u_0.$$

Esto sólo significa que $u = Me^{Jt}M^{-1}u_0$, y que la S y la Λ en la fórmula anterior $Se^{\Lambda t}S^{-1}u_0$ se han sustituido por M y J.

Conjunto de problemas B

1. Encuentre las formas de Jordan (¡en tres pasos!) de

$$A = \begin{bmatrix}1&1\\1&1\end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix}0&1&2\\0&0&0\\0&0&0\end{bmatrix}.$$

2. Demuestre que la solución especial u_2 en la ecuación (17) satisface $du/dt = Au$, exactamente debido a la cadena $Ax_1 = 8x_1$, $Ax_2 = 8x_2 + x_1$.

3. Para la matriz B en el problema 1, use $Me^{Jt}M^{-1}$ para calcular la exponencial e^{Bt}, y compárela con la serie de potencias $I + Bt + (Bt)^2/2! + \cdots$.

4. Demuestre que cada bloque de Jordan J_i es semejante a su traspuesta, $J_i^{\mathrm{T}} = P^{-1}J_iP$, usando la matriz permutación P con 1s a lo largo de la diagonal cruzada (de la parte inferior izquierda a la parte superior derecha). Concluya que toda matriz es semejante a su traspuesta.

5. "Por inspección", encuentre las formas de Jordan de

$$A = \begin{bmatrix}1&2&3\\0&4&5\\0&0&6\end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix}1&1\\-1&-1\end{bmatrix}.$$

6. Encuentre la forma de Jordan J y la matriz M para A y B (Los valores característicos de B son 1, 1, 1, −1). ¿Cuál es la solución de $du/dt = Au$, y cuál es e^{At}?

$$A = \begin{bmatrix}0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix}1&-1&0&-1\\0&2&0&1\\-2&1&-1&1\\2&-1&2&0\end{bmatrix}.$$

7. Suponga que $A^2 = A$. Demuestre que su forma de Jordan $J = M^{-1}AM$ satisface $J^2 = J$. Ya que los bloques diagonales permanecen separados, esto significa que $J_i^2 = J_i$ para cada bloque; demuestre por cálculo directo que J_i sólo puede ser un bloque de 1 por 1, $J_i = [0]$ o $J_i = [1]$. Así, A es semejante a una matriz diagonal de 0s y 1s.

Nota Este es un caso típico del teorema con que terminamos: *la matriz A puede diagonalizarse si y sólo si el producto $(A - \lambda_1 I)(A - \lambda_2 I)\cdots(A - \lambda_p I)$, sin incluir ninguna repetición de los λs, es cero.* Un caso extremo es una matriz con valores característicos distintos; el teorema de Cayley-Hamilton establece que con n factores, $A - \lambda I$ siempre se obtiene cero. El otro extremo es la matriz identidad, también diagonalizable ($p = 1$ y $A - I = 0$). La matriz no diagonalizable $A = \begin{bmatrix}1&1\\0&1\end{bmatrix}$ no satisface $(A - I) = 0$, sino sólo $(A - I = 0)^2$: una ecuación con una raíz repetida.



Soluciones a ejercicios seleccionados

Conjunto de problemas 1.2, página 9

1. Las rectas se cortan en $(x, y) = (3, 1)$. Así, 3(columna 1) + 1(columna 2) = (4, 4).

3. Estos "planos" se cortan en una recta en el espacio tetradimensional. El cuarto plano normalmente corta esta recta en un punto. Una ecuación inconsistente como $u + w = 5$ no deja solución (no hay intersección).

5. Los dos puntos sobre el plano son (1, 0, 0, 0) y (0, 1, 0, 0).

7. Resoluble para (3, 5, 8) y (1, 2, 3); no resoluble para $b = (3, 5, 7)$ o $b = (1, 2, 2)$.

9. Columna 3 = 2(columna 2) − columna 1. Si $b = (0, 0, 0)$, entonces $(u, v, w) = (c, -2c, c)$.

11. Tanto $a = 2$ y $a = -2$ proporcionan una recta de soluciones. Todas las demás a proporcionan $x = 0, y = 0$.

13. La representación de los renglones tiene dos rectas que se encuentran en (4, 2). La representación de las columnas tiene $4(1, 1) + 2(-2, 1) = 4$(columna 1) + 2(columna 2) = miembro derecho de (0, 6).

15. La representación de los renglones muestra cuatro *rectas*. La representación de las columnas está en el espacio *tetra*dimensional. No hay solución a menos que el miembro derecho sea una combinación de *las dos columnas*.

17. Si x, y, z satisfacen las dos primeras ecuaciones, también satisfacen la tercera ecuación. La recta **L** de soluciones contiene a $v = (1, 1, 0)$, $w = (\frac{1}{2}, 1, \frac{1}{2})$ y $u = \frac{1}{2}v + \frac{1}{2}w$, así como a todas las combinaciones $cv + dw$ con $c + d = 1$.

19. Columna 3 = columna 1; soluciones $(x, y, z) = (1, 1, 0)$ o $(0, 1, 1)$ y es posible sumar cualquier múltiplo de $(-1, 0, 1)$; $b = (4, 6, c)$ necesita $c = 10$ para resolubilidad.

21. Se cambian el segundo plano, el renglón 2 y todas las columnas de la matriz. La solución no cambia.

23. $u = 0, v = 0, w = 1$, porque 1(columna 3) = b.

Conjunto de problemas 1.3, página 15

1. Se multiplica por $\ell = \frac{10}{2} = 5$ y se resta para encontrar $2x + 3y = 1$ y $-6y = 6$. Los pivotes son 2 y −6

3. Se resta $-\frac{1}{2}$ veces la ecuación 1 (o se suma $\frac{1}{2}$ veces la ecuación 1). La nueva segunda ecuación es $3y = 3$. Así, $y = 1$ y $x = 5$. Si el miembro derecho cambia de signo, también lo hace la solución: $(x, y) = (-5, -1)$.

5. $6x + 4y$ es 2 veces $3x + 2y$. No hay solución, a menos que el miembro derecho sea $2 \cdot 10 = 20$. Por tanto, todos los puntos sobre la recta $3x + 2y = 10$ son soluciones, incluyendo a (0, 5) y (4, −1).

7. Si $a = 2$, la eliminación debe fallar. Las ecuaciones no tienen solución. Si $a = 0$, la eliminación se detiene para un intercambio de renglones. Así, con $3y = -3$ se obtiene $y = -1$ y con $4x + 6y = 6$ se obtiene $x = 3$.

9. $6x - 4y$ es 2 veces $(3x - 2y)$. Por consiguiente, se necesita $b_2 = 2b_1$. Entonces hay una infinidad de soluciones. Las columnas (3, 6) y (−2, −4) están sobre la misma recta.

11.

$2x - 3y = 3$	pro-	$2x - 3y = 3$		$x = 3$	Del renglón 2 se resta 2 × renglón 1
$y + z = 1$	por-	$y + z = 1$	y	$y = 1$	Del renglón 3 se resta 1 × renglón 1
$2y - 3z = 2$	ciona	$-5z = 0$		$z = 0$	Del renglón 3 se resta 2 × renglón 2

13. La segunda posición pivote contiene a $-2 - b$. Si $b = -2$, se intercambia con el renglón 3. Si $b = -1$ (caso singular), entonces la segunda ecuación es $-y - z = 0$. Una solución es $(1, 1, -1)$.

15. Si renglón 1 = renglón 2, entonces el renglón 2 es cero después del primer paso; se intercambia el renglón cero con el renglón 3 y no hay *tercer* pivote. Si columna 1 = columna 2, no hay *segundo* pivote.

17. El renglón 2 se convierte en $3y - 4z = 5$; luego, el renglón 3 se convierte en $(q + 4)z = t - 5$. Si $q = -4$, el sistema es singular; no hay tercer pivote. Luego, si $t = 5$, la tercera ecuación es $0 = 0$. Al escoger $z = 1$, la ecuación $3y - 4z = 5$ proporciona $y = 3$ y la ecuación 1 proporciona $x = -9$.

19. El sistema es singular si el renglón 3 es *una combinación de los renglones* 1 y 2. Desde la vista extrema, los tres planos forman un triángulo. Esto ocurre si los renglones 1 + 2 = renglón 3 en el miembro izquierdo pero no en el miembro derecho; por ejemplo, $x + y + z = 0$, $x - 2y - z = 1$, $2x - y = 9$. Ninguna pareja de planos es paralela, aunque sigue no habiendo solución.

21. El quinto pivote es $\frac{6}{5}$. El n-ésimo pivote es $\frac{(n+1)}{n}$.

23. Sistema triangular $\begin{cases} u + v + w = 2 \\ \quad 2v + 2w = -2 \\ \qquad 2w = 2 \end{cases}$ Solución $\begin{matrix} u = 3 \\ v = -2 \\ w = 1 \end{matrix}$.

25. $(u, v, w) = (3/2, 1/2, -3)$. Se cambia a +1 para que el sistema sea singular (2 columnas iguales).

27. $a = 0$ requiere un intercambio de renglones, pero el sistema es no singular: $a = 2$ lo hace singular (un pivote, infinidad de soluciones); $a = -2$ lo hace singular (un pivote, no hay solución).

29. El segundo término $bc + ad$ es $(a + b)(c + d) - ac - bd$ (sólo una multiplicación adicional).

31. La eliminación falla para $a = 2$ (columnas iguales), $a = 4$ (renglones iguales), $a = 0$ (columna de ceros).

Conjunto de problemas 1.4, página 26

1. $\begin{bmatrix} 17 \\ 4 \\ 17 \end{bmatrix}, \begin{bmatrix} 5 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}$. Con lados (2, 1) y (0, 3), el paralelogramo va a (2, 4).

3. Productos internos 54 y 0, al multiplicar la columna por el renglón se obtiene $\begin{bmatrix} 3 & 5 & 1 \\ -6 & -10 & -2 \\ 21 & 35 & 7 \end{bmatrix}$.

5. $Ax = (0, 0, 0)$, de modo que $x = (2, 1, 1)$ es una solución; las otras soluciones son $cx = (2c, c, c)$.

7. Ejemplos: Diagonal $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 7 \end{bmatrix}$, simétrica $\begin{bmatrix} 1 & 3 & 4 \\ 3 & 2 & 0 \\ 4 & 0 & 7 \end{bmatrix}$, triangular $\begin{bmatrix} 1 & 3 & 4 \\ 0 & 2 & 0 \\ 0 & 0 & 7 \end{bmatrix}$, simétrica sesgada $\begin{bmatrix} 0 & 3 & 4 \\ -3 & 0 & 0 \\ -4 & 0 & 0 \end{bmatrix}$.

9. *a)* a_{11} *b)* $\ell_{i1} = a_{i1}/a_{11}$ *c)* la nueva a_{ij} es $a_{ij} - \frac{a_{i1}}{a_{11}} a_{1j}$

b) segundo pivote $a_{22} - \frac{a_{21}}{a_{11}} a_{12}$.

11. Los coeficientes del renglón B son 2, 1, 4 provenientes de A. El primer renglón de AB es $[6 \;\; 3]$.

13. $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, D = A, E = F = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$.

15. $AB_1 = B_1A$ da $b = c = 0$. $AB_2 = B_2A$ da $a = d$. Así $A = aI$.

17. $A(A+B) + B(A+B)$, $(A+B)(B+A)$, $A^2 + AB + BA + B^2$ siempre es igual a $(A+B)^2$.

19. $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} p & q \\ r & s \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix} [p \;\; q] + \begin{bmatrix} b \\ d \end{bmatrix} [r \;\; s] = \begin{bmatrix} ap+br & aq+bs \\ cp+dr & cq+ds \end{bmatrix}$.

21. $A^n = A$; $B^n = \begin{bmatrix} 1 & 0 \\ 0 & (-1)^n \end{bmatrix}, C = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ = matriz cero.

23. $E_{32}E_{21}b = (1, -5, -35)$ pero $E_{21}E_{32}b = (1, -5, \mathbf{0})$. Así, el renglón 3 no siente ningún efecto debido al renglón 1.

25. Al cambiar a_{33} de 7 a 11 se cambia el tercer pivote de 5 a 9. Al cambiar a_{33} de 7 a 2 se cambia el pivote de 5 a *ningún* pivote.

27. Para invertir E_{31}, se *suma* 7 veces el renglón 1 al renglón 3.

La matriz es $R_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 7 & 0 & 1 \end{bmatrix}$.

29. $E_{13} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}; E_{31}E_{13} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$. ¡Pruebe con la matriz identidad!

31. E_{21} tiene $\ell_{21} = -\frac{1}{2}$, E_{32} tiene $\ell_{32} = -\frac{2}{3}$, E_{43} tiene $\ell_{43} = -\frac{3}{4}$. En caso contrario, las Es coinciden con I.

33. $\begin{aligned} a + b + c &= 4 \\ a + 2b + 4c &= 8 \\ a + 3b + 9c &= 14 \end{aligned}$ proporciona $\begin{aligned} a &= 2 \\ b &= 1. \\ c &= 1 \end{aligned}$

35. *a*) Cada columna es E multiplicado por una columna de B.

b) $EB = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 1 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \end{bmatrix}$.

Los renglones de EB son combinaciones de los renglones de B, por lo que son múltiplos de $[1 \;\; 2 \;\; 4]$.

37. (renglón 3) $\cdot x$ es $\sum a_{3j}x_j$, y $(A^2)_{11}$ = (renglón 1) $\cdot$ (columna 1) $= \sum a_{1j}a_{j1}$.

39. $BA = 3I$ es de 5 por 5, $AB = 5I$ es de 3 por 3, $ABD = 5D$ es de 3 por 1, ABD: No, $A(B + C)$: No

41. *a*) $B = 4I$. *b*) $B = 0$. *c*) $B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.

d) Todo renglón de B es 1, 0, 0.

43. *a*) mn (todo) *b*) mnp. *c*) n^3 (esto es n^2 productos punto)

45. $\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} [3 \;\; 3 \;\; 0] + \begin{bmatrix} 0 \\ 4 \\ 1 \end{bmatrix} [1 \;\; 2 \;\; 1] = \begin{bmatrix} 3 & 3 & 0 \\ 6 & 6 & 0 \\ 6 & 6 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 4 & 8 & 4 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 0 \\ 10 & 14 & 4 \\ 7 & 8 & 1 \end{bmatrix}$.

47. A multiplicada por B es $A\begin{bmatrix} \vert & \vert & \vert \end{bmatrix}, \begin{bmatrix} \text{———} \\ \text{———} \\ \text{———} \end{bmatrix}B, \begin{bmatrix} \text{———} \\ \text{———} \\ \text{———} \end{bmatrix}\begin{bmatrix} \vert & \vert & \vert \end{bmatrix}, \begin{bmatrix} \vert & \vert & \vert \end{bmatrix}\begin{bmatrix} \text{———} \\ \text{———} \\ \text{———} \end{bmatrix}$.

49. Los bloques (2, 2) $S = D - CA^{-1}B$ constituyen el **complemento de Schur**: los bloques en $d - (cb/a)$.

51. A multiplicada por $X = [x_1 \;\; x_2 \;\; x_3]$ es la matriz identidad $I = [Ax_1 \;\; Ax_2 \;\; Ax_3]$.

53. $\begin{bmatrix} a+b & a+b \\ c+d & c+d \end{bmatrix}$ coincide con $\begin{bmatrix} a+c & b+d \\ a+c & b+d \end{bmatrix}$ cuando $b = c$ y $a = d$.

55. $2x + 3y + z + 5t = 8$ es $Ax = b$ con la matriz $A = [2 \;\; 3 \;\; 1 \;\; 5]$ de 1 por 4. Las soluciones x llenan un "plano" tridimensional en cuatro dimensiones.

57. El producto punto $[1 \;\; 4 \;\; 5] \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ = (1 por 3)(3 por 1) es cero para puntos (x, y, z) sobre un plano $x + 4y + 5z = 0$ en tres dimensiones. Las columnas de A son vectores unidimensionales.

59. A * v = [3 4 5]′ y v′ * v = 50; v * A proporcionan un mensaje de error.

61. $M = \begin{bmatrix} 8 & 3 & 4 \\ 1 & 5 & 9 \\ 6 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 5+u & 5-u+v & 5-v \\ 5-u-v & 5 & 5+u+v \\ 5+v & 5+u-v & 5-u \end{bmatrix}$;

$M_3(1,1,1) = (15, 15, 15)$; $M_4(1,1,1,1) = (34, 34, 34, 34)$ porque la suma de los números 1 a 16 es 136, que es 4(34).

Conjunto de problemas 1.5, página 39

1. U es no singular cuando ningún elemento sobre la diagonal principal es cero.

3. $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} = I$ también.

$(E^{-1}F^{-1}G^{-1})(GFE) = E^{-1}F^{-1}FE = E^{-1}E = I$; también $(GFE)(E^{-1}F^{-1}G^{-1}) = I$.

5. $LU = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 & 3 \\ 0 & 5 & 7 \\ 0 & 0 & -1 \end{bmatrix}$; después de la eliminación, $\begin{bmatrix} 2 & 3 & 3 \\ 0 & 5 & 7 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$.

7. $FGH = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{bmatrix}; HGF = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 4 & 2 & 1 & 0 \\ 8 & 4 & 2 & 1 \end{bmatrix}$.

9. a) No singular cuando $d_1d_2d_3 \neq 0$. b) Suponga que $d_3 \neq 0$. $Lc = b$ se resuelve yendo hacia abajo: $Lc = b$ proporciona $c = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Así, $\begin{bmatrix} d_1 & -d_1 & 0 \\ 0 & d_2 & -d_2 \\ 0 & 0 & d_3 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

proporciona $x = \begin{bmatrix} 1/d_3 \\ 1/d_3 \\ 1/d_3 \end{bmatrix}$.

11. Al resolver $Lc = b$ yendo hacia abajo se obtiene $c = \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix}$; Al resolver $Ux = c$ yendo hacia arriba se obtiene $x = \begin{bmatrix} 5 \\ -2 \\ 0 \end{bmatrix}$.

13. Permutación renglones 2 y 3 $\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}, \begin{bmatrix}u\\v\\w\end{bmatrix} = \begin{bmatrix}2\\-3\\4\end{bmatrix};$

Permutación renglones 1 y 2 $\begin{bmatrix}0&1&0\\1&0&0\\0&0&0\end{bmatrix}, \begin{bmatrix}u\\v\\w\end{bmatrix} = \begin{bmatrix}1\\-1\\1\end{bmatrix}.$

15. $PA = LDU$ es $\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}\begin{bmatrix}0&1&1\\1&0&1\\2&3&4\end{bmatrix} = \begin{bmatrix}1&0&0\\0&1&0\\2&3&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\0&0&-1\end{bmatrix}\begin{bmatrix}1&0&1\\0&1&1\\0&0&1\end{bmatrix};$

$PA = LDU$ es $\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}\begin{bmatrix}1&2&1\\2&4&2\\1&1&1\end{bmatrix} = \begin{bmatrix}1&0&0\\1&1&0\\2&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&-1&0\\0&0&0\end{bmatrix}\begin{bmatrix}1&2&1\\0&1&0\\0&0&0\end{bmatrix}.$

17. L se convierte en $\begin{bmatrix}1&0&0\\1&1&0\\2&0&1\end{bmatrix}$. MATLAB y otros códigos utilizan $PA = LU$.

19. $a = 4$ conduce a un intercambio de renglones; $3b + 10a = 40$ conduce a una matriz singular; $c = 0$ conduce a un intercambio de renglones; $c = 3$ conduce a una matriz singular.

21. $\ell_{31} = 1$ y $\ell_{32} = 2$ ($\ell_{33} = 1$): invierta los pasos para recuperar $x + 3y + 6z = 11$ a partir de $Ux = c$; 1 multiplicado por $(x + y + z = 5)$ + 2 multiplicado por $(y + 2z = 2)$ + 1 multiplicado por $(z = 2)$ proporciona $x + 3y + 6z = 11$.

23. $\begin{bmatrix}1&&\\0&1&\\0&-2&1\end{bmatrix}\begin{bmatrix}1&&\\-2&1&\\0&0&1\end{bmatrix}A = \begin{bmatrix}1&1&1\\0&2&3\\0&0&-6\end{bmatrix} = U.$

$A = \begin{bmatrix}1&0&0\\2&1&0\\0&2&1\end{bmatrix}U = E_{21}^{-1}E_{32}^{-1}U = LU.$

25. De 2 por 2: $d = 0$ no está permitido.

$\begin{bmatrix}1&1&0\\1&1&2\\1&2&1\end{bmatrix} = \begin{bmatrix}1&&\\\ell&1&\\m&n&1\end{bmatrix}\begin{bmatrix}d&e&g\\&f&h\\&&i\end{bmatrix}$ $d = 1$, $e = 1$, entonces $\ell = 1$; $f = 0$ no está permitido **sin pivote en el renglón 2.**

27. $A = \begin{bmatrix}2&4&8\\0&3&9\\0&0&7\end{bmatrix}$ tiene $L = I$ y $D = \begin{bmatrix}2&&\\&3&\\&&7\end{bmatrix}$; $A = LU$ tiene $U = A$ (pivotes en la diagonal); $A = LDU$ tiene $U = D^{-1}A = \begin{bmatrix}1&2&4\\0&1&3\\0&0&1\end{bmatrix}$ con unos en la diagonal.

29. $\begin{bmatrix}a&a&a&a\\a&b&b&b\\a&b&c&c\\a&b&c&d\end{bmatrix} = \begin{bmatrix}1&&&\\1&1&&\\1&1&1&\\1&1&1&1\end{bmatrix}\begin{bmatrix}a&a&a&a\\&b-a&b-a&b-a\\&&c-b&c-b\\&&&d-c\end{bmatrix}$. Requiere $a \neq 0$, $b \neq a$, $c \neq b$, $d \neq c$.

31. $\begin{bmatrix}1&&\\1&1&\\0&1&1\end{bmatrix}\begin{bmatrix}1&1&0\\&1&1\\&&1\end{bmatrix} = LIU;$ $\begin{bmatrix}a&a&0\\a&a+b&b\\0&b&b+c\end{bmatrix}$ = (misma L) $\begin{bmatrix}a&&\\&b&\\&&c\end{bmatrix}$ (misma U)

33. $\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}c = \begin{bmatrix}4\\5\\6\end{bmatrix}$ proporciona $c = \begin{bmatrix}4\\1\\1\end{bmatrix}$. $\begin{bmatrix}1&1&1\\0&1&1\\0&0&1\end{bmatrix}x = \begin{bmatrix}4\\1\\1\end{bmatrix}$ proporciona $x = \begin{bmatrix}3\\0\\1\end{bmatrix}$. $A = LU$.

35. La submatriz superior de 2 por 2 contiene los dos primeros pivotes 2 y 7; Razón: La eliminación sobre A comienza en la esquina superior izquierda con eliminación sobre B.

37. $\begin{bmatrix}1&1&1&1&1\\1&2&3&4&5\\1&3&6&10&15\\1&4&10&20&35\\1&5&15&35&70\end{bmatrix}$

$= \begin{bmatrix}1&&&&\\1&1&&&\\1&2&1&&\\1&3&3&1&\\1&4&6&4&1\end{bmatrix}\begin{bmatrix}1&1&1&1&1\\&1&2&3&4\\&&1&3&6\\&&&1&4\\&&&&1\end{bmatrix}.$ *El triángulo de Pascal* en L y U. El código lu de MATLAB arruina el patrón. El código chol no realiza ningún intercambio de renglones para matrices simétricas con pivotes positivos.

39. Cada nuevo *miembro derecho* sólo cuesta n^2 pasos en comparación con los $n^3/3$ para eliminación total $A\backslash b$.

41. 2 intercambios; 3 intercambios; 50 intercambios y luego 51.

43. $P = \begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}$; $P_1 = \begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}$ y $P_2 = \begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}$

(P_2 proporciona un intercambio de columnas).

45. Hay $n!$ matrices permutación de orden n. Al final, dos potencias de P *deben ser las mismas*: Si $P^r = P^s$, entonces $P^{r-s} = I$. Ciertamente, $r - s \leq n!$

$P = \begin{bmatrix}P_2&\\&P_3\end{bmatrix}$ es 5 por 5 con $P_2 = \begin{bmatrix}0&1\\1&0\end{bmatrix}$, $P_3 = \begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}$, y $P^6 = I$.

47. La solución es $x = (1, 1, \ldots, 1)$. Así, $x = Px$.

Conjunto de problemas 1.6, página 52

1. $A_1^{-1} = \begin{bmatrix}0&\frac{1}{3}\\\frac{1}{2}&0\end{bmatrix}$; $A_2^{-1} = \begin{bmatrix}\frac{1}{2}&0\\-1&\frac{1}{2}\end{bmatrix}$; $A_3^{-1} = \begin{bmatrix}\cos\theta&\operatorname{sen}\theta\\-\operatorname{sen}\theta&\cos\theta\end{bmatrix}$.

3. $A^{-1} = BC^{-1}$; $A^{-1} = U^{-1}L^{-1}P$

5. $A(AB)$ = (quitar paréntesis) = $(A^2)(B) = I$.

7. $\begin{bmatrix}\sqrt{3}/2 & 1/2\\ 1/2 & -\sqrt{3}/2\end{bmatrix}, \begin{bmatrix}-\sqrt{3}/2 & 1/2\\ 1/2 & \sqrt{3}/2\end{bmatrix}, \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$ todas tienen $A^2 = I$.

9. Si el renglón 3 de A^{-1} fuese (a, b, c, d), entonces con $A^{-1}A = I$ se obtendría $2a = 0$, $a + 3b = 0$, $4a + 8b = 1$. Esto no tiene solución.

11. a) $\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix} = \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$. b) $\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} + \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$.

c) $\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix} = \begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}$; $(B^{-1} + A^{-1})^{-1} = B(A + B)^{-1}A$.

13. $A^TB = 8$; $B^TA = 8$; $AB^T = \begin{bmatrix}6 & 6\\ 2 & 2\end{bmatrix}$; $BA^T = \begin{bmatrix}6 & 2\\ 6 & 2\end{bmatrix}$.

15. a) $n(n + 1)/2$ elementos sobre y arriba de la diagonal. b) $(n - 1)n/2$ elementos arriba de la diagonal.

17. a) La inversa de una matriz triangular inferior (superior) sigue siendo triangular inferior (superior). Al multiplicar matrices triangulares inferiores (superiores) se obtiene una matriz triangular inferior (superior). b) Las diagonales principales de $L_1^{-1}L_2D_2 = L_2D_2$ y $D_1U_1U_2^{-1}$ son las mismas que las de D_2 y D_1, respectivamente. $L_1^{-1}L_2D_2 = D_1U_1U_2^{-1}$, de modo que se tiene $D_1 = D_2$. Al comparar los elementos fuera de las diagonales de $L_1^{-1}L_2D_2 = D_1U_1U_2^{-1}$, ambas matrices deben ser diagonales. $L_1^{-1}L_2D_2 = D_2$, $D_1U_1U_2^{-1} = D_1$, D_1 es invertible, de modo que $L_1^{-1}L_2 = I$, $U_1U_2^{-1} = I$. Así, $L_1 = L_2$, $U_1 = U_2$.

19. $\begin{bmatrix}1 & 0 & 0\\ 3 & 1 & 0\\ 5 & 1 & 1\end{bmatrix} \begin{bmatrix}1 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & 2\end{bmatrix} \begin{bmatrix}1 & 3 & 5\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}$;

$\begin{bmatrix}1 & 0\\ b/a & 1\end{bmatrix} \begin{bmatrix}a & 0\\ 0 & d - (b^2/a)\end{bmatrix} \begin{bmatrix}1 & b/a\\ 0 & 1\end{bmatrix} = LDL^T$.

21. A partir de $B(I - AB) = (I - BA)B$ se obtiene $(I - BA)^{-1} = B(I - AB)^{-1}B^{-1}$, una inversa explícita en el supuesto de que B e $I - AB$ sean invertibles. *Segundo método*: si $I - AB$ no es invertible, entonces $BAx = x$ para alguna x diferente de cero. En consecuencia, $ABAx = Ax$, o bien, $ABy = y$, e $I - AB$ podría no ser invertible. (Observe que $y = Ax$ es diferente de cero, con base en $BAx = x$.)

23. $\begin{bmatrix}x\\ y\end{bmatrix} = \begin{bmatrix}0.5\\ -0.2\end{bmatrix}$, $\begin{bmatrix}t\\ z\end{bmatrix} = \begin{bmatrix}-0.2\\ 0.1\end{bmatrix}$ de modo que $A^{-1} = \frac{1}{10}\begin{bmatrix}5 & -2\\ -2 & 1\end{bmatrix}$.

25. a) En $Ax = (1, 0, 0)$, ecuación 1 + ecuación 2 − ecuación 3 es $0 = 1$. b) Los miembros derechos deben satisfacer $b_1 + b_2 = b_3$. c) El renglón 3 se convierte en un renglón de ceros; no hay tercer pivote.

27. Si B intercambia los renglones 1 y 2 de A, entonces B^{-1} intercambia las *columnas* 1 y 2 de A^{-1}.

29. Si A tiene una columna de ceros, también BA. Así, $BA = I$ es imposible. No existe A^{-1}.

31. $\begin{bmatrix}1 & & \\ & 1 & \\ & -1 & 1\end{bmatrix} \begin{bmatrix}1 & & \\ & 1 & \\ -1 & & 1\end{bmatrix} \begin{bmatrix}1 & & \\ -1 & 1 & \\ & & 1\end{bmatrix} = \begin{bmatrix}1 & & \\ -1 & 1 & \\ 0 & -1 & 1\end{bmatrix} = E$;

entonces $\begin{bmatrix}1 & & \\ 1 & 1 & \\ 1 & 1 & 1\end{bmatrix}$ es $L = E^{-1}$, después de invertir el orden de estas tres matrices elementales y cambiar −1 a +1.

33. A ∗ ones(4,1) proporciona el vector cero, de modo que A no puede ser invertible.

35. $\begin{bmatrix}1 & 3 & 1 & 0\\ 2 & 7 & 0 & 1\end{bmatrix} \to \begin{bmatrix}1 & 3 & 1 & 0\\ 0 & 1 & -2 & 1\end{bmatrix} \to \begin{bmatrix}1 & 0 & 7 & -3\\ 0 & 1 & -2 & 1\end{bmatrix} = [I \quad A^{-1}]$;

$\begin{bmatrix}1 & 3 & 1 & 0\\ 3 & 8 & 0 & 1\end{bmatrix} \to \begin{bmatrix}1 & 0 & -8 & 3\\ 0 & 1 & 3 & -1\end{bmatrix} = [I \quad A^{-1}]$.

37. $\begin{bmatrix}1 & a & b & 1 & 0 & 0\\ 0 & 1 & c & 0 & 1 & 0\\ 0 & 0 & 1 & 0 & 0 & 1\end{bmatrix} \to \begin{bmatrix}1 & a & 0 & 1 & 0 & -b\\ 0 & 1 & 0 & 0 & 1 & -c\\ 0 & 0 & 1 & 0 & 0 & 1\end{bmatrix}$

$\to \begin{bmatrix}1 & 0 & 0 & 1 & -a & ac - b\\ 0 & 1 & 0 & 0 & 1 & -c\\ 0 & 0 & 1 & 0 & 0 & 1\end{bmatrix}$.

39. $\begin{bmatrix}2 & 2 & 0 & 1\\ 0 & 2 & 1 & 0\end{bmatrix} \to \begin{bmatrix}2 & 0 & -1 & 1\\ 0 & 2 & 1 & 0\end{bmatrix} \to \begin{bmatrix}1 & 0 & -1/2 & 1/2\\ 0 & 1 & 1/2 & 0\end{bmatrix} = [I \quad A^{-1}]$.

41. No es invertible para $c = 7$ (columnas iguales), $c = 2$ (renglones iguales), $c = 0$ (columna cero).

43. $A^{-1} = \begin{bmatrix}1 & 1 & 0 & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 1\end{bmatrix}$. La matriz A^{-1} de 5 por 5 tiene 1s en la diagonal y arriba de ésta.

45. $\begin{bmatrix}I & 0\\ -C & I\end{bmatrix}$, $\begin{bmatrix}A^{-1} & 0\\ -D^{-1}CA^{-1} & D^{-1}\end{bmatrix}$, y $\begin{bmatrix}-D & I\\ I & 0\end{bmatrix}$.

47. Para $Ax = b$ con A = ones(4, 4) = matriz singular y b = ones(4, 1), A\b, escoge $x = (1, 0, 0, 0)$ y pinv(A) ∗ b escoge la solución más corta $x = (1, 1, 1, 1)/4$.

49. $A^T = \begin{bmatrix}1 & 9\\ 0 & 3\end{bmatrix}$, $A^{-1} = \begin{bmatrix}1 & 0\\ -3 & 1/3\end{bmatrix}$, $(A^{-1})^T = (A^T)^{-1} = \begin{bmatrix}1 & -3\\ 0 & 1/3\end{bmatrix}$; $A^T = A$ y

entonces $A^{-1} = \frac{1}{c^2}\begin{bmatrix}0 & c\\ c & -1\end{bmatrix} = (A^{-1})^T = (A^T)^{-1}$.

51. $\left((AB)^{-1}\right)^T = (B^{-1}A^{-1})^T = (A^{-1})^T(B^{-1})^T$; $(U^{-1})^T$ es triangular *inferior*.

53. a) $x^TAy = a_{22} = 5$. b) $x^TA = [4 \quad 5 \quad 6]$. c) $Ay = \begin{bmatrix}2\\ 5\end{bmatrix}$.

55. $(Px)^T(Py) = x^TP^TPy = x^Ty$ ya que $P^TP = I$; por lo general, $Px \cdot y = x \cdot P^Ty \neq x \cdot Py$:

$$\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{bmatrix}\begin{bmatrix}1\\ 2\\ 3\end{bmatrix} \cdot \begin{bmatrix}1\\ 1\\ 2\end{bmatrix} \neq \begin{bmatrix}1\\ 2\\ 3\end{bmatrix} \cdot \begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{bmatrix}\begin{bmatrix}1\\ 1\\ 2\end{bmatrix}.$$

57. PAP^T recupera la simetría.

59. a) La traspuesta de R^TAR es $R^TA^TR^{TT} = R^TAR = n$ por n.
b) $(R^TR)_{jj}$ = (columna j de R) · (columna j de R) = longitud al cuadrado de la columna j.

61. Las corrientes totales son $A^{\mathrm{T}}y = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & -1 \end{bmatrix} \begin{bmatrix} y_{BC} \\ y_{CS} \\ y_{BS} \end{bmatrix} = \begin{bmatrix} y_{BC} + y_{BS} \\ -y_{BC} + y_{CS} \\ -y_{CS} - y_{BS} \end{bmatrix}$.

De cualquier forma $(Ax)^{\mathrm{T}}y = x^{\mathrm{T}}(A^{\mathrm{T}}y) = x_B y_{BC} + x_B y_{BS} - x_C y_{BC} + x_C y_{CS} - x_S y_{CS} - x_S y_{BS}$.

63. $Ax \cdot y$ es el *costo* de las entradas, mientras $x \cdot A^{\mathrm{T}}y$ es el *valor* de las salidas.

65. Estos son grupos: triangular inferior con diagonal de unos, diagonal invertible D y las permutaciones P. Dos más: permutaciones pares; todas las matrices no singulares.

67. Al reordenar los renglones y/o las columnas de $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ se mueve el elemento a, con lo que no se obtiene $\begin{bmatrix} a & c \\ b & d \end{bmatrix}$.

69. Casi seguramente, las matrices aleatorias son invertibles.

71. La matriz $-1, 2, -1$ en la sección 1.7 tiene $A = LDL^{\mathrm{T}}$ con $\ell_{i,i-1} = 1 - \frac{1}{i}$.

Conjunto de problemas 1.7, página 63

1. $$\begin{bmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ -\frac{1}{2} & 1 & & \\ & -\frac{2}{3} & 1 & \\ & & -\frac{3}{4} & 1 \end{bmatrix} \begin{bmatrix} 2 & & & \\ & \frac{3}{2} & & \\ & & \frac{4}{3} & \\ & & & \frac{5}{4} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & & \\ & 1 & -\frac{2}{3} & \\ & & 1 & -\frac{3}{4} \\ & & & 1 \end{bmatrix} = LDL^{\mathrm{T}}$$
det = 5.

3. $A_0 = \begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{bmatrix}$. La suma de cada renglón es 1, por lo que $A_0 \begin{bmatrix} c \\ c \\ c \\ c \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$.

5. $(u_1, u_2, u_3) = (\pi^2/8, 0, -\pi^2/8)$ en vez de los valores verdaderos $(1, 0, -1)$.

7. $H^{-1} = \begin{bmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{bmatrix}$.

9. La matriz de Hilbert de 10 por 10 está muy mal condicionada.

11. Un pivote grande se multiplica por menos que 1 en la eliminación de cada elemento abajo del pivote. Un caso extremo, con multiplicadores = 1 y pivotes = $\frac{1}{2}, \frac{1}{2}, 4$, es

$$A = \begin{bmatrix} 1/2 & 1/2 & 1 \\ -1/2 & 0 & 1 \\ -1/2 & -1 & 1 \end{bmatrix}.$$

Conjunto de problemas 2.1, página 73

1. a) El conjunto de todas las (u, v), donde u y v son razones p/q de enteros. b) El conjunto de todas las (u, v), donde $u = 0$ o $v = 0$.

3. $\mathbf{C}(A)$ es el eje x; $\mathbf{N}(A)$ es la recta que pasa por $(1, 1)$; $\mathbf{C}(B)$ es $\mathbf{R}^2$; $\mathbf{N}(B)$ es la recta que pasa por $(-2, 1, 0)$; $\mathbf{C}(C)$ es el punto $(0, 0)$ en $\mathbf{R}^2$; el espacio nulo $\mathbf{N}(C)$ es $\mathbf{R}^3$.

5. Reglas que no se cumplen: a) 7, 8, b) 1 c) 1, 2, 8.

7. b), d) y e) son subespacios. No es posible multiplicar por -1 en a) y c). No es posible sumar en f).

9. La suma de dos matrices no singulares puede ser singular $(A + (-A))$. La suma de dos matrices singulares puede ser no singular.

11. a) Una posibilidad: las matrices cA forman un subespacio que no contiene a B.
b) Sí; el subespacio debe contener a $A - B = I$.
c) El subespacio de matrices cuya diagonal principal es toda cero.

13. Si $(f + g)(x)$ es la $f(g(x))$ de costumbre, entonces $(g + f)(x)$ es $g(f(x))$, que es diferente. En la regla 2 ambos miembros son $f(g(h(x)))$. La regla 4 se rompe porque podría no existir función inversa $f^{-1}(x)$ tal que $f(f^{-1}(x)) = x$. Si esta función inversa existe, se trata del vector $-f$.

15. La suma de $(4, 0, 0)$ y $(0, 4, 0)$ no está sobre el plano; tiene $x + y - 2z = 8$.

17. a) Los subespacios de $\mathbf{R}^2$ son $\mathbf{R}^2$ mismo, las rectas que pasan por $(0, 0)$ y el punto $(0, 0)$.
b) Los subespacios de $\mathbf{R}^4$ son $\mathbf{R}^4$ mismo, los planos tridimensionales $n \cdot v = 0$, los subespacios bidimensionales ($n_1 \cdot v = 0$ y $n_2 \cdot v = 0$), las rectas unidimensionales que sólo pasan por $(0, 0, 0, 0)$ y $(0, 0, 0, 0)$.

19. El menor subespacio que contiene a $\mathbf{P}$ y $\mathbf{L}$ es $\mathbf{P}$ o $\mathbf{R}^3$.

21. El espacio columna de A es el eje x = todos los vectores $(x, 0, 0)$. El espacio columna de B es el plano x-y = todos los vectores $(x, y, 0)$. El espacio columna de C es la recta de los vectores $(x, 2x, 0)$.

23. Una combinación de las columnas de C también es una combinación de las columnas de A (mismo espacio columna; B tiene un espacio columna diferente).

25. La columna adicional b agranda el espacio columna, a menos que b *ya esté* en ese espacio:

$[A \;\; b] = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ (espacio de columna más grande) (no hay solución para $Ax = b$).

$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ (b incluida en el espacio de columna) ($Ax = b$ tiene solución).

27. Espacio columna = $\mathbf{R}^8$. Todo b es una combinación de las columnas, ya que $Ax = b$ es resoluble.

29. $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ o bien $\begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$; $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 3 & 6 & 0 \end{bmatrix}$ (columnas en la línea 1).

31. $\mathbf{R}^2$ contiene vectores con *dos* componentes: no pertenecen a $\mathbf{R}^3$.

Conjunto de problemas 2.2, página 85

1. $x + y + z = 1, x + y + z = 0$. Al cambiar 1 a 0, $(x, y, z) = c(-1, 1, 0) + d(-1, 0, 1)$.

3. Forma escalonada $U = \begin{bmatrix} 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$; variables libres x_1, x_3, x_4; soluciones especiales $(1, 0, 0, 0)$, $(0, 0, 1, 0)$ y $(0, -3, 0, 1)$. Consistente cuando $b_2 = 2b_1$. Solución completa $(0, b_1, 0, 0)$ más cualquier combinación de soluciones especiales.

5. $\begin{bmatrix}u\\v\\w\end{bmatrix} = \begin{bmatrix}-2v-3\\v\\2\end{bmatrix} = v\begin{bmatrix}-2\\1\\0\end{bmatrix} + \begin{bmatrix}-3\\0\\2\end{bmatrix}$; ¡No hay solución!

7. $c = 7$ permite $u = 1$, $v = 1$, $w = 0$. El espacio columna es un plano.

9. a) $x = x_2\begin{bmatrix}-2\\1\\0\\0\end{bmatrix} + x_4\begin{bmatrix}2\\0\\-2\\1\end{bmatrix}$, para cualesquiera x_2, x_4. Reducida por renglones $R = \begin{bmatrix}1&2&0&-2\\0&0&1&2\\0&0&0&0\end{bmatrix}$.

b) Solución completa $x = \begin{bmatrix}a-3b\\0\\b\\0\end{bmatrix} + x_2\begin{bmatrix}-2\\1\\0\\0\end{bmatrix} + x_4\begin{bmatrix}2\\0\\-2\\1\end{bmatrix}$, para cualesquiera x_2, x_4

11. $\begin{bmatrix}1&1\\1&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}$ tiene espacio nulo = recta que pasa por $(-1, 1)$, pero no tiene solución. Cualquier $b = \begin{bmatrix}c\\c\end{bmatrix}$ tiene muchas soluciones particulares de $Ax_p = b$.

13. $R = \begin{bmatrix}1&1&1&1\\0&0&0&0\\0&0&0&0\end{bmatrix}$; $R = \begin{bmatrix}1&0&1&0\\0&1&0&1\\0&0&0&0\\0&0&0&0\end{bmatrix}$; $R = \begin{bmatrix}1&-1&1&-1\\0&0&0&0\\0&0&0&0\end{bmatrix}$. (a) $r = 1$. (b) $r = 2$. (c) $r = 1$.

15. Una matriz espacio nulo $N = \begin{bmatrix}-F\\I\end{bmatrix}$ es de n por $n - r$.

17. Creo que es cierto.

19. Las soluciones especiales son las columnas de $N = \begin{bmatrix}-2&-3\\-4&-5\\1&0\\0&1\end{bmatrix}$ y $N = \begin{bmatrix}1&0\\0&-2\\0&1\end{bmatrix}$.

21. Las r columnas pivote de A forman una submatriz de m por r de rango r, de modo que la matriz A^* tiene r renglones pivote independientes, con lo cual se obtiene una submatriz invertible de r por r de A. (Los renglones pivote de A^* y A son los mismos, ya que la eliminación se efectúa en el mismo orden: simplemente para A^* no vemos las columnas "libres" de ceros que aparecen para A.)

23. $(uv^T)(wz^T) = u(v^Tw)z^T$ tiene rango 1, a menos que $v^Tw = 0$.

25. Se tiene que $AB = I$, cuyo rango es n. Luego, rango$(AB) \leq$ rango(A) obliga a que rango$(A) = n$.

27. Si $R = EA$ y la misma $R = E^*B$, entonces $B = (E^*)^{-1}EA$. (Para obtener B, A se reduce a R y luego se invierten los pasos de vuelta a B). B es una matriz *invertible* multiplicada por A, cuando comparten la misma R.

29. Debido a que R empieza con r renglones independientes, R^T empieza con r columnas independientes (y por tanto ceros). Así, *su* forma escalonada reducida es $\begin{bmatrix}I&0\\0&0\end{bmatrix}$, donde I es de r por r.

31. Si $c = 1$, $R = \begin{bmatrix}1&1&2&2\\0&0&0&0\\0&0&0&0\end{bmatrix}$ tiene x_2, x_3, x_4 libres.

Si $c \neq 1$, $R = \begin{bmatrix}1&0&2&2\\0&1&0&0\\0&0&0&0\end{bmatrix}$ tiene x_3, x_4 libres.

Soluciones especiales en $N = \begin{bmatrix}-1&-2&-2\\1&0&0\\0&1&0\\0&0&1\end{bmatrix}$ $(c = 1)$ y $N = \begin{bmatrix}-2&-2\\0&0\\1&0\\0&1\end{bmatrix}$ $(c \neq 1)$.

Si $c = 1$, $R = \begin{bmatrix}0&1\\0&0\end{bmatrix}$ tiene x_1 libres; si $c = 2$, $R = \begin{bmatrix}1&-2\\0&0\end{bmatrix}$ tiene x_2 libres; $R = I$ si $c \neq 1, 2$.

Soluciones especiales en $N = \begin{bmatrix}1\\0\end{bmatrix}$ $(c = 1)$ o bien, $N = \begin{bmatrix}2\\1\end{bmatrix}$ $(c = 2)$ o bien, $N =$ *matriz vacía* de 2 por 0.

33. $x_{\text{completa}} = \begin{bmatrix}-2\\0\\1\end{bmatrix} + x_2\begin{bmatrix}-3\\1\\0\end{bmatrix}$; $x_{\text{completa}} = \begin{bmatrix}1/2\\0\\1/2\\0\end{bmatrix} + x_2\begin{bmatrix}-3\\1\\0\\0\end{bmatrix} + x_4\begin{bmatrix}0\\0\\-2\\1\end{bmatrix}$.

35. a) Resoluble si $b_2 = 2b_1$ y $3b_1 - 3b_3 + b_4 = 0$. Así, $x = \begin{bmatrix}5b_1 - 2b_3\\b_3 - 2b_1\end{bmatrix}$ (no hay variables libres). b) Resoluble si $b_2 = 2b_1$ y $3b_1 - 3b_3 + b_4 = 0$.
Así, $x = \begin{bmatrix}5b_1 - 2b_3\\b_3 - 2b_1\\0\end{bmatrix} + x_3\begin{bmatrix}-1\\-1\\1\end{bmatrix}$.

37. Un sistema de 1 por 3 tiene por lo menos dos variables libres.

39. a) La solución particular x_p siempre se multiplica por 1. b) Cualquier solución puede ser x_p. c) $\begin{bmatrix}3&3\\3&3\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}6\\6\end{bmatrix}$. Así, $\begin{bmatrix}1\\1\end{bmatrix}$ es más corta (longitud $\sqrt{2}$) que $\begin{bmatrix}2\\0\end{bmatrix}$. d) La solución "homogénea" en el espacio nulo es $x_n = 0$ cuando A es invertible.

41. x_p se multiplica por 2, misma x_n; $\begin{bmatrix}x\\X\end{bmatrix}_p$ es $\begin{bmatrix}x_p\\0\end{bmatrix}$; las soluciones especiales también incluyen a las columnas de $\begin{bmatrix}-I\\I\end{bmatrix}$; x_p y las soluciones especiales no cambian.

43. Para A, $q = 3$ proporciona rango 1, y cualquier otra q proporciona rango 2. Para B, $q = 6$ proporciona rango 1, y cualquier otra q proporciona rango 2.

45. a) $r < m$, siempre $r \leq n$. b) $r = m$, $r < n$. c) $r < m$, $r = n$. d) $r = m = n$.

47. $R = \begin{bmatrix}1&0&0&\mathbf{0}\\0&0&1&\mathbf{0}\\0&0&0&\mathbf{0}\end{bmatrix}$ y $x_n = \begin{bmatrix}0\\1\\0\end{bmatrix}$; $\begin{bmatrix}1&0&0&\mathbf{-1}\\0&0&1&\mathbf{2}\\0&0&0&\mathbf{5}\end{bmatrix}$: no hay solución debido al renglón 3.

49. $A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 0 & 3 \end{bmatrix}$; B no puede existir porque dos ecuaciones en tres incógnitas no pueden tener una solución.

51. El rango de A es $4 - 1 = 3$; la solución completa de $Ax = 0$ es $x = (2, 3, 1, 0)$.

$$R = \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{con } -2,\ -3 \text{ en la columna libre.}$$

53. a) Falso. b) Verdadero. c) Verdadero (sólo n columnas). d) Verdadero (sólo m renglones).

55. $$U = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{y} \quad R = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$ (R no proviene de esta U).

57. Si columna 1 = columna 5, entonces x_5 es una variable libre. Su solución especial es $(-1, 0, 0, 0, 1)$.

59. Es seguro que la columna 5 no tiene pivote, ya que es una combinación de columnas anteriores y x_5 es libre. Con cuatro pivotes en las otras columnas, la solución especial es $(1, 0, 1, 0, 1)$. El espacio nulo contiene a todos los múltiplos de $(1, 0, 1, 0, 1)$ (una recta en $\mathbf{R}^5$).

61. $A = \begin{bmatrix} 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & -2 \end{bmatrix}$.

63. Esta construcción es imposible; dos columnas pivote, dos variables libres, sólo tres columnas.

65. $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

67. Lo más probable es que R sea I; lo más probable es que R sea I con el cuarto renglón de ceros.

69. Cualquier renglón cero aparece después de estos renglones: $R = [1 \ \ -2 \ \ -3]$, $R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $R = I$

Conjunto de problemas 2.3, página 98

1. $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = 0$ proporciona $c_3 = c_2 = c_1 = 0$. Pero $v_1 + v_2 - 4v_3 + v_4 = 0$ (dependiente).

3. Si $a = 0$, entonces columna 1 = 0; si $d = 0$, entonces b(columna 1) $- a$(columna 2) $= 0$; si $f = 0$, entonces todas las columnas terminan en cero (todas son perpendiculares a $(0, 0, 1)$, y todas en el plano xy, deben ser dependientes).

5. a) $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & -5 & -7 \\ 0 & -1 & -5 \end{bmatrix}$

$\to \begin{bmatrix} 1 & 2 & 3 \\ 0 & -5 & -7 \\ 0 & 0 & -18/5 \end{bmatrix}$: invertible $\Rightarrow$ columnas independientes (hubieran podido usarse renglones).

b) $\begin{bmatrix} 1 & 2 & -3 \\ -3 & 1 & 2 \\ 2 & -3 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -3 \\ 0 & 7 & -7 \\ 0 & 0 & 0 \end{bmatrix}$; $A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, la suma de las columnas es 0 (hubieran podido usarse renglones).

7. La suma $v_1 - v_2 + v_3 = 0$ porque $(w_2 - w_3) - (w_1 - w_3) + (w_1 - w_2) = 0$.

9. a) Los cuatro vectores son las columnas de una matriz A de 3 por 4 con por lo menos una variable libre, de modo que $Ax = 0$. b) Dependientes si el rango de $[v_1\ v_2]$ es 0 o 1. c) $0v_1 + c(0, 0, 0) = 0$ tiene una solución distinta de cero (tome cualquier $c \neq 0$).

11. a) Recta en $\mathbf{R}^3$. b) Plano en $\mathbf{R}^3$. c) Plano en $\mathbf{R}^3$. d) Todo $\mathbf{R}^3$.

13. Todas las dimensiones son 2. Los espacios renglón de A y U son los mismos.

15. $v = \frac{1}{2}(v + w) + \frac{1}{2}(v - w)$ y $w = \frac{1}{2}(v + w) - \frac{1}{2}(v - w)$. Los dos pares *generan* el mismo espacio. Son una base cuando v y w son *independientes*.

17. Si la eliminación produce uno o más renglones cero, entonces los renglones de A son linealmente dependientes;

por ejemplo, en el problema 16 $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

$\to \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

19. Los n vectores independientes generan un espacio de dimensión n. Forman una *base* para ese espacio. Si son las columnas de A, entonces m *no es menor* que n ($m \geq n$).

21. $C(U)$: Bases cualesquiera de $\mathbf{R}^2$; $N(U)$: (renglón 1 y renglón 2) o bien, (renglón 1 y renglón 1 + renglón 2).

23. Columnas independientes $\Rightarrow$ rango n. Las columnas generan $\mathbf{R}^m \Rightarrow$ rango m. Las columnas son una base para $\mathbf{R}^m \Rightarrow$ rango $= m = n$.

25. a) La única solución es $x = 0$ porque *las columnas son independientes*. b) $Ax = b$ es resoluble porque *las columnas generan* $\mathbf{R}^5$.

27. Las columnas 1 y 2 son bases para los espacios columna (diferentes) de A y U; los renglones 1 y 2 son bases para los espacios renglón (iguales); $(1, -1, 1)$ es una base para los espacios nulos (iguales).

29. rango$(A) = 2$ si $c = 0$ y $d = 2$; rango $(B) = 2$, excepto cuando $c = d$ o $c = -d$.

31. Sean $v_1 = (1, 0, 0, 0), \ldots, v_4 = (0, 0, 0, 1)$ los vectores de coordenadas. Si $\mathbf{W}$ es la recta que pasa por $(1, 2, 3, 4)$, ninguno de los vs está en $\mathbf{W}$

33. a) Si no hubiera una base, sería posible sumar más vectores independientes, con lo cual se excedería la dimensión k. b) Si no hubiera una base, sería posible eliminar algunos vectores, dejando menos que la dimensión k.

35. a) Falso, podría no haber solución. b) verdadero, 7 vectores en $\mathbf{R}^5$ son dependientes.

37. a) $\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix}$, $\begin{bmatrix}0&0&0\\0&1&0\\0&0&0\end{bmatrix}$, $\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}$. b) Sumar $\begin{bmatrix}0&1&0\\1&0&0\\0&0&0\end{bmatrix}$, $\begin{bmatrix}0&0&1\\0&0&0\\1&0&0\end{bmatrix}$, $\begin{bmatrix}0&0&0\\0&0&1\\0&1&0\end{bmatrix}$. c) $\begin{bmatrix}0&1&0\\-1&0&0\\0&0&0\end{bmatrix}$, $\begin{bmatrix}0&0&1\\0&0&0\\-1&0&0\end{bmatrix}$, $\begin{bmatrix}0&0&0\\0&0&1\\0&-1&0\end{bmatrix}$ son una base para todas las $A = -A^{\mathrm{T}}$.

39. $y(0) = 0$ requiere $A + B + C = 0$. Una base es $\cos x - \cos 2x$ y $\cos x - \cos 3x$.

41. $y_1(x), y_2(x), y_3(x)$ pueden ser $x, 2x, 3x$ (dim 1) o bien, $x, 2x, x^2$ (dim 2), o bien, x, x^2, x^3 (dim 3).

43.
$$\begin{bmatrix}1&&\\&1&\\&&1\end{bmatrix} = \begin{bmatrix}&1&\\1&&\\&&1\end{bmatrix} - \begin{bmatrix}&1&\\&&1\\1&&\end{bmatrix} + \begin{bmatrix}&&1\\&1&\\1&&\end{bmatrix} + \begin{bmatrix}1&&\\&&1\\&1&\end{bmatrix} - \begin{bmatrix}&&1\\1&&\\&1&\end{bmatrix}.$$

Compruebe el elemento (1, 1), luego (2, 3), luego (3, 3), luego (1, 2) para demostrar que estos cinco Ps son independientes. Cuatro condiciones sobre los nueve elementos hacen que las sumas de renglones y columnas sean iguales: suma del renglón 1 = suma del renglón 2 = suma del renglón 3 = suma de la columna 1 = suma de la columna 2 (= suma de la columna 3 es automático porque suma de todos los renglones = suma de todas las columnas).

45. Si la matriz $[A\ b]$ de 5 por 5 es invertible, b no es una combinación de las columnas de A. Si $[A\ b]$ es singular y las columnas de A son independientes, b es una combinación de estas columnas.

Conjunto de problemas 2.4, página 110

1. Falso; lo único que se sabe es que las dimensiones son iguales. El espacio nulo izquierdo tiene una dimensión más pequeña dim $= m - r$.

3. $C(A)$: $r = 2, (1, 0, 1)\ (0, 1, 0)$; $N(A)$: $n - r = 2, (2, -1, 1, 0), (-1, 0, 0, 1)$;
$C(A^{\mathrm{T}})$: $r = 2, (1, 2, 0, 1), (0, 1, 1, 0)$; $N(A^{\mathrm{T}})$: $m - r = 1, (-1, 0, 1)$;
$C(U)$: $(1, 0, 0), (0, 1, 0)$; $N(U)$: $(2, -1, 1, 0), (-1, 0, 0, 0)$;
$C(U^{\mathrm{T}})$: $(1, 2, 0, 1), (0, 1, 1, 0)$; $N(A^{\mathrm{T}})$: $(0, 0, 1)$.

5. A multiplicada por cada columna de B es cero, de modo que $C(B)$ está contenido en el espacio nulo $N(A)$.

7. Con base en $Ax = 0$, el espacio renglón y el espacio nulo deben ser ortogonales. Consulte el capítulo 3.

9. $[1\ \ 2\ \ 4]$; $\begin{bmatrix}1&2&4\\2&4&8\\3&6&12\end{bmatrix}$ tiene el mismo espacio nulo.

11. Si $Ax = 0$ tiene una solución diferente de cero, entonces $r < n$ y $C(A^{\mathrm{T}})$ es menor que $\mathbf{R}^n$. Así, $A^{\mathrm{T}}y = f$ no es resoluble para alguna f. Ejemplo: $A = [1\ \ 1]$ y $f = (1, 2)$.

13. $d = bc/a$; el único pivote es a.

15. Con columnas independientes: rango n; espacio nulo $= \{0\}$; el espacio renglón es $\mathbf{R}^n$; inversa *izquierda*.

17. $A = [1\ \ 1\ \ 0]$; $B = [0\ \ 0\ \ 1]$.

19. No: por ejemplo, todas las matrices invertibles de n por n tienen los mismos cuatro subespacios.

21. a) $\begin{bmatrix}1&0\\1&0\\0&1\end{bmatrix}$. b) Imposible: dimensiones $1 + 1 \neq 3$ c) $[1\ \ 1]$.

d) $\begin{bmatrix}-9&-3\\3&1\end{bmatrix}$. e) Imposible: espacio renglón = espacio columna requiere $m = n$. Así, $m - r = n - r$.

23. A invertible: base del espacio renglón = base del espacio columna = (1, 0, 0), (0, 1, 0), (0, 0, 1); las bases del espacio nulo y del espacio nulo izquierdo son vacías. B: base del espacio renglón (1, 0, 0, 1, 0, 0), (0, 1, 0, 0, 1, 0) y (0, 0, 1, 0, 0, 1); base del espacio columna (1, 0, 0), (0, 1, 0), (0, 0, 1); base del espacio nulo $(-1, 0, 0, 1, 0, 0)$, $(0, -1, 0, 0, 1, 0)$ y $(0, 0, -1, 0, 0, 1)$; la base del espacio nulo izquierdo es vacía.

25. a) El espacio renglón es igual al espacio nulo. En consecuencia, el rango (dimensión del espacio renglón) es el mismo. b) El espacio columna es igual al espacio nulo izquierdo. Mismo rango (dimensión del espacio columna).

27. a) La inexistencia de una solución significa que $r < m$. Siempre se tiene $r \leq n$. No es posible comparar m y n.
b) Si $m - r > 0$, el espacio nulo de A^{T} contiene un vector diferente de cero.

29. Base del espacio renglón (1, 2, 3, 4), (0, 1, 2, 3), (0, 0, 1, 2); base del espacio nulo $(0, 1, -2, 1)$; base del espacio columna (1, 0, 0), (0, 1, 0), (0, 0, 1); el espacio nulo izquierdo tiene una base vacía.

31. Si $Av = 0$ y v es un renglón de A, entonces $v \cdot v = 0$. Sólo $v = 0$ en ambos espacios.

33. Renglón 3 − 2(renglón 2) + renglón 1 = renglón cero, de modo que los vectores $c(1, -2, 1)$ están en el espacio nulo izquierdo. Ocurre que los mismos vectores están en el espacio nulo.

35. a) u y w generan a $C(A)$. b) v y z generan a $C(A^{\mathrm{T}})$. c) Rango < 2 si u y w son dependientes o v y z son dependientes. d) El rango de $uv^{\mathrm{T}} + wz^{\mathrm{T}}$ es 2.

37. a) Verdadero (mismo rango). b) Falso ($A = [1\ \ 0]$. c) Falso (A puede ser invertible y también no simétrica). d) Verdadero.

39. $a_{11} = 1, a_{12} = 0, a_{13} = 1, a_{22} = 0, a_{32} = 1, a_{31} = 0, a_{23} = 1, a_{33} = 0, a_{21} = 1$ (no es única).

41. Rango $r = n$ significa espacio nulo = vector cero y $x_n = 0$.

Conjunto de problemas 2.5, página 122

1. $A = \begin{bmatrix}1&-1&0\\0&1&-1\\1&0&-1\end{bmatrix}$; $N(A)$ contiene múltiplos de $\begin{bmatrix}1\\1\\1\end{bmatrix}$; $N(A^{\mathrm{T}})$ contiene múltiplos de $\begin{bmatrix}1\\1\\-1\end{bmatrix}$.

3. La suma de los elementos en cada renglón es cero. Por tanto, cualquier combinación posee esa misma propiedad: $f_1 + f_2 + f_3 = 0$; $A^{\mathrm{T}}y = f \Rightarrow y_1 + y_3 = f_1$, $-y_1 + y_2 = f_2$, $-y_2 - y_3 = f_3 \Rightarrow f_1 + f_2 + f_3 = 0$. Significa que la corriente total que entra desde el exterior es cero.

5. $\begin{bmatrix} c_1+c_3 & -c_1 & -c_3 \\ -c_1 & c_1+c_2 & -c_2 \\ -c_3 & -c_2 & c_2+c_3 \end{bmatrix}$; $\begin{bmatrix} c_1+c_3 & -c_1 \\ -c_1 & c_1+c_2 \end{bmatrix}$ tiene pivotes $c_1 + c_3$, $\dfrac{c_1c_3 + c_1c_2 + c_2c_3}{c_1 + c_3}$.

7. Las condiciones sobre b son $b_1 + b_4 - b_5 = 0$, $b_3 - b_4 + b_6 = 0$, $b_2 - b_5 + b_6 = 0$.

9. $\begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix}$, $\begin{bmatrix} c_1+c_2+c_5 & -c_1 & -c_2 & -c_5 \\ -c_1 & c_1+c_3+c_4 & -c_3 & -c_4 \\ -c_2 & -c_3 & c_2+c_3+c_6 & -c_6 \\ -c_5 & -c_4 & -c_6 & c_4+c_5+c_6 \end{bmatrix}$.
Estas cs que se unen con el nodo j ahora aparecen en el renglón j.

11. $\begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ f_1 \\ f_2 \\ f_3 \end{bmatrix}$; $x = \begin{bmatrix} -4 \\ -\frac{5}{3} \\ -\frac{14}{3} \\ 0 \end{bmatrix}$; $y = \begin{bmatrix} -\frac{7}{3} \\ \frac{4}{3} \\ \frac{10}{3} \\ \frac{14}{3} \end{bmatrix}$.

13. Hay 20 elecciones de 3 aristas de 6, ya que "6 en 3" $= \dfrac{6!}{3!3!} = 20$. Cuatro elecciones proporcionan triángulos, dejando 16 árboles generadores.

15. Creo que ya está considerada.

17. 9 nodos − 12 aristas + 4 ciclos = 1; 7 nodos − 12 aristas + 6 ciclos = 1.

19. Con $x = (1, 1, 1, 1)$ se obtiene $Ax = 0$; así, $A^TAx = 0$; de nuevo, el rango es $n - 1$.

21. $M = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$ y

$M^2 = \begin{bmatrix} 3 & 2 & 2 & 2 \\ 2 & 3 & 2 & 2 \\ 2 & 2 & 3 & 2 \\ 2 & 2 & 2 & 3 \end{bmatrix}$. $(M^2)_{ij} = a_{i1}a_{1j} + \cdots + a_{in}a_{nj}$ y se obtiene $a_{ik}a_{kj} = 1$ cuando hay una ruta de 2 pasos que va de i a k a j. Observe las tres rutas que van de un nodo a sí mismo.

Conjunto de problemas 2.6, página 133

1. Rotación $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$; $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

3. $\|Ax\|^2 = 1$ siempre produce una elipse.

5. Son transformados a $(1, 3)$, $(2, 6)$, $(-1, -3)$. El eje x gira; las rectas verticales se desplazan hacia arriba o hacia abajo, pero permanecen verticales.

7. Matriz de segundas derivadas $\begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. El espacio nulo es generado por $(1, 0, 0, 0)$ y $(0, 1, 0, 0)$, que proporcionan una P_1 lineal. Las segundas derivadas de funciones lineales son cero. Accidentalmente, el espacio columna es el mismo que el espacio nulo, porque las segundas derivadas de cúbicas son lineales.

9. e^t y e^{-t} son una base para las soluciones de $u'' = u$.

11. $\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $\quad H^2 = I$.

13. a) Sí. b) Sí. ¡Para $(AB)C$ o $A(BC)$ no se requieren paréntesis!

15. $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ y $A^2 = I$; la doble traspuesta de una matriz proporciona la matriz misma. Observe que $A_{23} = 1$ porque la traspuesta de la matriz 2 es la matriz 3.

17. $A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; $B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$; $AB = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$; $BA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

19. a) es invertible con $T^{-1}(y) = y^{1/3}$; c) es invertible con $T^{-1}(y) = y - 11$.

21. Con $w = 0$, la linealidad proporciona $T(v + 0) = T(v) + T(0)$. Así, $T(0) = 0$. Con $c = -1$, la linealidad proporciona $T(-0) = -T(0)$. Con certeza $T(-0) = T(0)$. Así, $T(0) = 0$.

23. $S(T(v)) = S(v) = v$.

25. b) y c) son lineales, a) falla $T(2v) = 2T(v)$, d) falla $T(v + w) = T(v) + T(w)$.

27. $T(T(v)) = (v_3, v_1, v_2)$; $T^3(v) = v$; $T^{100}(v) = T(T^{99}(v)) = T(v)$.

29. a) $T(1, 0) = 0$. b) $(0, 0, 1)$ no está en el rango. c) $T(0, 1) = 0$.

31. La ley asociativa proporciona $A(M_1 + M_2) = AM_1 + AM_2$. La ley distributiva aplicada a las cs proporciona $A(cM) = c(AM)$.

33. Con ninguna matriz se obtiene $A\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Para los profesores: la dimensión del espacio de matrices es 4. Las transformaciones lineales en ese espacio deben provenir de matrices de 4 por 4 (16 parámetros). Las multiplicaciones por A en los problemas 31 y 32 son transformaciones especiales con sólo 4 parámetros.

35. $T(I) = 0$ pero $M = \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix} = T(M)$; éstas llenan el rango. $M = \begin{bmatrix} a & 0 \\ c & d \end{bmatrix}$ en el núcleo.

37. a) $M = \begin{bmatrix} r & s \\ t & u \end{bmatrix}$. b) $N = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1}$. c) $ad = bc$.

39. Reordenar la base con la *matriz permutación*; cambiar las longitudes por medio de la *matriz diagonal positiva*.

41. $\begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$; el determinante de VanderMonde $= (b - a)(c - a)(c - b)$; los puntos a, b y c deben ser distintos, y así el determinante $\neq 0$ (es posible interpolar).

43. Si T no es invertible, entonces $T(v_1), \ldots, T(v_n)$ no es una base. Así, no es posible escoger $w_i = T(v_i)$ como base de salida.

45. $S(T(v)) = (-1, 2)$, pero $S(v) = (-2, 1)$ y $T(S(v)) = (1, -2)$. Así, $TS \neq ST$.

47. La matriz H de Hadamard tiene columnas ortogonales de longitud 2. Por tanto, la inversa de H es $H^T/4 = H/4$.

49. Falso: los n vectores diferentes de cero deben ser independientes.

Conjunto de problemas 3.1, página 148

1. $\|x\| = \sqrt{21}$; $\|y\| = 3\sqrt{2}$; $x^Ty = 0$.

3. $(x_2/x_1)(y_2/y_1) = -1$ significa que $x_1y_1 + x_2y_2 = 0$, de modo que $x^Ty = 0$.

5. v_1 y v_3 son ortogonales, así como v_2 y v_3.

7. $x = (-2, 1, 0)$; $y = (-1, -1, 1)$; el renglón $z = (1, 2, 1)$ es ortogonal al espacio nulo.

9. El complemento ortogonal es la recta que pasa por $(-1, -1, 1)$ y $(0, 0, 0)$.

11. Si $A^Ty = 0$, entonces $y^Tb = y^TAx = (y^TA)x = 0$, lo cual contradice que $y^Tb \neq 0$.

13. La figura separa cualquier y en $\mathbf{R}^m$ en la parte del espacio columna + la parte del espacio nulo izquierdo.

15. No puede existir una matriz así, ya que $(1, 2, 1)^T(1, -2, 1) \neq 0$.

17. La matriz con la base para $\mathbf{V}$ como sus renglones. Así, el espacio nulo es $\mathbf{V}^\perp = \mathbf{W}$.

19. a) Si $\mathbf{V}$ y $\mathbf{W}$ son rectas en $\mathbf{R}^3$, $\mathbf{V}^\perp$ y $\mathbf{W}^\perp$ son planos que se cortan. b) $\mathbf{V}$.

21. $(1, 2, -1)$ es perpendicular a $\mathbf{P}$. $A = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 2 \end{bmatrix}$ tiene $N(A) = \mathbf{P}$; $B = [1 \quad 2 \quad -1]$ tiene espacio renglón $= \mathbf{P}$.

23. $A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$ tiene subespacios = cuatro rectas; $(1, 1)$ ortogonal a $(-1, 1)$, $(1, 2)$ ortogonal a $(-2, 1)$. El espacio renglón siempre es perpendicular al espacio nulo $\perp$.

25. a) $\begin{bmatrix} 1 & 2 & -3 \\ 2 & -3 & 1 \\ -3 & 5 & -2 \end{bmatrix}$. b) $\begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$ no es ortogonal a $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. c) $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ en $C(A)$ y $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ en $N(A^T)$ es imposible: no son perpendiculares. d) $A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$ tiene $A^2 = 0$.
e) $(1, 1, 1)$ está en el espacio nulo y en el espacio renglón; no existe una matriz así.

27. a) Si $Ax = b$ tiene una solución y $A^Ty = 0$, entonces $b^Ty = (Ax)^Ty = 0$. b) b no está en el espacio columna; así, no es perpendicular a ninguna y en el espacio nulo izquierdo.

29. $x = x_r + x_n$, donde x_r está en el espacio renglón y x_n está en el espacio nulo. Así, $Ax_n = 0$ y $Ax = Ax_r + Ax_n = Ax_r$. Todos los vectores Ax son combinaciones de las columnas de A. Si $x = (1, 0)$, entonces $x_r = (1/2, 1/2)$.

31. a) Para una matriz simétrica, el espacio columna y el espacio renglón son iguales.
b) x está en el espacio nulo y z está en el espacio columna = espacio renglón; por tanto, estos "vectores característicos" tienen $x^Tz = 0$.

33. x se separa en $x_r + x_n = (1, -1) + (1, 1) = (2, 0)$.

35. $Ax = B\hat{x}$ significa que $[\,A \quad B\,]\begin{bmatrix} x \\ -\hat{x} \end{bmatrix} = 0$. Tres ecuaciones homogéneas en cuatro incógnitas siempre tienen una solución diferente de cero. Aquí $x = (3, 1)$ y $\hat{x} = (1, 0)$, y $Ax = B\hat{x} = (5, 6, 5)$ están en ambos espacios columna. ¡Dos planos en $\mathbf{R}^3$ (que pasan por cero) deben cortarse por lo menos en una recta!

37. $A^Ty = 0$ proporciona $(Ax)^Ty = x^TA^Ty = 0$. Así, $y \perp Ax$ y $N(A^T) \perp C(A)$.

39. $\mathbf{S}^\perp$ es el espacio nulo de $A = \begin{bmatrix} 1 & 5 & 1 \\ 2 & 2 & 2 \end{bmatrix}$. En consecuencia, $\mathbf{S}^\perp$ es un *subespacio* incluso si $\mathbf{S}$ no lo es.

41. Si $\mathbf{V}$ es todo $\mathbf{R}^4$, entonces $\mathbf{V}^\perp$ sólo contiene al *vector cero*. Así, $(\mathbf{V}^\perp)^\perp = \mathbf{R}^4 = \mathbf{V}$.

43. $(1, 1, 1, 1)$ es una base de $\mathbf{P}^\perp$. El espacio nulo de $A = [1 \; 1 \; 1 \; 1]$ es el plano $\mathbf{P}$.

45. La columna 1 de A^{-1} es ortogonal al espacio generado por los renglones 2o., . . . , n-ésimo de A.

47. $A = \begin{bmatrix} 2 & 2 & -1 \\ -1 & 2 & 2 \\ 2 & -1 & 2 \end{bmatrix}$,
$A^TA = 9I$ es *diagonal*: $(A^TA)_{ij}$ = (columna i de A) · (columna j).

49. a) $(1, -1, 0)$ está en ambos planos. Los vectores normales son perpendiculares, ¡y los planos siguen cortándose! b) Se requieren *tres* vectores ortogonales para generar todo el complemento ortogonal en $\mathbf{R}^5$. c) Las rectas pueden cortarse sin ser ortogonales.

51. Cuando $AB = 0$, el espacio columna de B está contenido en el espacio nulo de A. En consecuencia, dimensión de $C(B) \leq$ dimensión de $N(A)$. Esto quiere decir rango $(B) \leq 4 -$ rango(A).

Conjunto de problemas 3.2, página 157

1. a) $(x + y)/2 \geq \sqrt{xy}$ (media aritmética $\geq$ media geométrica de x y y).
b) $\|x + y\|^2 \leq (\|x\| + \|y\|)^2$ significa que $(x+y)^T(x+y) \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2$. El miembro izquierdo es $x^Tx + 2x^Ty + y^Ty$. Después de cancelar lo anterior, se obtiene $x^Ty \leq \|x\|\|y\|$.

3. $p = (10/3, 10/3, 10/3)$; $(5/9, 10/9, 10/9)$.

5. $\cos\theta = 1/\sqrt{n}$, de modo que $\theta = \arccos(1/\sqrt{n})$; $P = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} [1/n \quad \cdots \quad 1/n] =$ todos los elementos $\dfrac{1}{n}$.

7. Se escoge $b = (1, \ldots, 1)$; la igualdad ocurre si $a_1 = \cdots = a_n$ (entonces a es paralela a b).

9. $P^2 = \dfrac{aa^Taa^T}{a^Taa^Ta} = \dfrac{a(a^Ta)a^T}{(a^Ta)(a^Ta)} = \dfrac{aa^T}{a^Ta} = P$.

11. a) $P = \begin{bmatrix} \frac{1}{10} & \frac{3}{10} \\ \frac{3}{10} & \frac{9}{10} \end{bmatrix}$; $P_2 = I - P_1 = \begin{bmatrix} \frac{9}{10} & -\frac{3}{10} \\ -\frac{3}{10} & \frac{1}{10} \end{bmatrix}$. b) $P_1 + P_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$;
$P_1P_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. La suma de las proyecciones sobre dos rectas perpendiculares proporciona el vector mismo. La proyección sobre una recta y luego una recta perpendicular proporcionan el vector cero.

13. Traza $= \dfrac{a_1a_1}{a^Ta} + \cdots + \dfrac{a_na_n}{a^Ta} = \dfrac{a^Ta}{a^Ta} = 1$.

15. $\|Ax\|^2 = (Ax)^T(Ax) = xA^TAx$, $\|A^Tx\|^2 = (A^Tx)^T(A^Tx) = xAA^Tx$. Si $A^TA = AA^T$, entonces $\|Ax\| = \|A^Tx\|$. (Estas matrices se denominan *normales*).

17. a) $a^Tb/a^Ta = 5/3$; $p = (5/3, 5/3, 5/3)$; $e = (-2/3, 1/3, 1/3)$ tiene $e^Ta = 0$. b) $a^Tb/a^Ta = -1$; $p = (1, 3, 1) = b$ y $e = (0, 0, 0)$.

19. $P_1 = \frac{1}{3}\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix} = P_1^2$ y $P_1b = \frac{1}{3}\begin{bmatrix}5\\5\\5\end{bmatrix}$. $P_2 = \frac{1}{11}\begin{bmatrix}1&3&1\\3&9&3\\1&3&1\end{bmatrix}$ y

$P_2b = \begin{bmatrix}1\\3\\1\end{bmatrix}$.

21. $P_1 = \frac{1}{9}\begin{bmatrix}1&-2&-2\\-2&4&4\\-2&4&4\end{bmatrix}$, $P_2 = \frac{1}{9}\begin{bmatrix}4&4&-2\\4&4&-2\\-2&-2&1\end{bmatrix}$.

$P_1P_2 =$ matriz cero porque $a_1 \perp a_2$.

23. $P_1 + P_2 + P_3$

$$= \frac{1}{9}\begin{bmatrix}1&-2&-2\\-2&4&4\\-2&4&4\end{bmatrix} + \frac{1}{9}\begin{bmatrix}4&4&-2\\4&4&-2\\-2&-2&1\end{bmatrix} + \frac{1}{9}\begin{bmatrix}4&-2&4\\-2&1&-2\\4&-2&4\end{bmatrix} = I.$$

25. Debido a que A es invertible, $P = A(A^TA)^{-1}A^T = AA^{-1}(A^T)^{-1}A^T = I$: proyectar sobre todo $\mathbf{R}^2$.

Conjunto de problemas 3.3, página 170

1. $\hat{x} = 2$; $E^2 = (10 - 3x)^2 + (5 - 4x)^2$ es minimizado; $(4, -3)^T(3, 4) = 0$.

3. $\hat{x} = \begin{bmatrix}\frac{1}{3}\\\frac{1}{3}\end{bmatrix}$; $p = \begin{bmatrix}\frac{1}{3}\\\frac{1}{3}\\\frac{2}{3}\end{bmatrix}$; $b - p = \begin{bmatrix}\frac{2}{3}\\\frac{2}{3}\\-\frac{2}{3}\end{bmatrix}$ es perpendicular a ambas columnas.

5. $b = 4, 5, 9$ en $t = -1, 0, 1$; la mejor recta es $6 + (5/2)t$; $p = (7/2, 6, 17/2)$.

7. $P = A(A^TA)^{-1}A^T = \begin{bmatrix}1&1/2&0\\1/2&1/2&-1/2\\0&-1/2&1\end{bmatrix}$.

9. a) $P^T = (P^TP)^T = P$. Así, $P = P^TP = P^2$. b) P proyecta sobre el espacio $\mathbf{Z} = \{0\}$.

11. $P + Q = I$, $PQ = 0$, trasponer a $QP = 0$, de modo que $(P - Q)(P - Q) = P - 0 - 0 + Q = I$.

13. La mejor recta es $61/35 - (36/35)t$; $p = (133/35, 95/35, 61/35, -11/35)$ a partir de $C = Dt$.

15. $H^2 = (I - 2P)^2 = I - 4P + 4P^2 = I - 4P + 4P = I$. Con dos reflexiones se obtiene I.

17. La proyección sobre $x + y = 0 =$ Proyección sobre $(-1, 1) = \begin{bmatrix}1/2&-1/2\\-1/2&1/2\end{bmatrix}$.

19. La matriz proyección sobre el espacio renglón debería ser $A^T(AA^T)^{-1}A$ si los renglones fuesen independientes.

21. La mejor recta es $\begin{bmatrix}a_1^Ta_1&-a_1^Ta_2\\-a_1^Ta_2&a_2^Ta_2\end{bmatrix}\begin{bmatrix}\hat{x}_1\\\hat{x}_2\end{bmatrix} = \begin{bmatrix}a_1^Tb\\-a_2^Tb\end{bmatrix}$; $\hat{x} = \begin{bmatrix}2\\1\end{bmatrix}$.

23. $C = (A^TA)^{-1}A^Tb$, $A^T = [1 \cdots 1]$, $b = (y_1, \ldots, y_m)^T$ entonces

$$C = \frac{A^Tb}{A^TA} = \frac{y_1 + \cdots + y_m}{m}.$$

25. $A = \begin{bmatrix}1&-1&1\\1&0&0\\1&1&1\\1&2&4\end{bmatrix}$; $x = \begin{bmatrix}C\\D\\E\end{bmatrix}$; $b = \begin{bmatrix}2\\0\\-3\\-5\end{bmatrix}$.

27. a) $a^Ta = m$, $a^Tb = b_1 + \cdots + b_m$. En consecuencia, $\hat{x}$ es la media de las bs. b) La varianza es $\|e\|^2 = \sum_{i=1}^m (b_i - \hat{x})^2$. (c) $p = (3, 3, 3)$, $e = (-2, -1, 3)$, $p^Te = 0$.

$P = \frac{1}{3}\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}$.

29. $(\hat{x} - x)(\hat{x} - x)^T = (A^TA)^{-1}A^T[(b - Ax)(b - Ax)^T]A(A^TA)^{-1}$. Para errores independientes, al sustituir $(b - Ax)(b - Ax)^T = \sigma^2 I$ se obtiene la *matriz de covarianza* $(A^TA)^{-1}A^T\sigma^2A(A^TA)^{-1}$. Esto se simplifica a $\sigma^2(A^TA)^{-1}$ fórmula clara para la matriz de covarianza.

31. $\frac{1}{10}b_{10} + \frac{9}{10}\hat{x}_9 = \frac{1}{10}(b_1 + \cdots + b_{10})$.

33. $\begin{bmatrix}1&0\\1&1\\1&3\\1&4\end{bmatrix}\begin{bmatrix}C\\D\end{bmatrix} = \begin{bmatrix}0\\8\\8\\20\end{bmatrix}$. Al cambiar el miembro derecho por $p = \begin{bmatrix}1\\5\\13\\17\end{bmatrix}$; $\hat{x} = \begin{bmatrix}1\\4\end{bmatrix}$ se resuelve $A\hat{x} = p$.

35. Parábola más próxima: $\begin{bmatrix}1&0&0\\1&1&1\\1&3&9\\1&4&16\end{bmatrix}\begin{bmatrix}C\\D\\E\end{bmatrix} = \begin{bmatrix}0\\8\\8\\20\end{bmatrix}$.

$$A^TA\hat{x} = \begin{bmatrix}4&8&26\\8&26&92\\26&92&338\end{bmatrix}\begin{bmatrix}\hat{C}\\\hat{D}\\\hat{E}\end{bmatrix} = \begin{bmatrix}36\\112\\400\end{bmatrix}.$$

37. a) La mejor recta es $x = 1 + 4t$, que pasa por el punto central $(\hat{t}, \hat{b}) = (2, 9)$. b) Con base en la primera ecuación $Cm + D\sum t_i = \sum b_i$. Se divide entre m para obtener $C + D\hat{t} = \hat{b}$.

39. $\hat{x}_W = \dfrac{w_1^2b_1 + \cdots + w_m^2b_m}{w_1^2 + \cdots + w_m^2}$.

41. $\hat{x}_W = (1/21, 4/7)$; $A\hat{x}_W = (1/21, 13/21, 25/21)$, $b - A\hat{x}_W = (-1/21, 8/21, -4/21)$, $(A\hat{x}_W)W^TW(b - A\hat{x}_W) = 0$.

Conjunto de problemas 3.4, página 185

1. a) $-4 = C - 2D$, $-3 = C - D$, $-1 = C + D$, $0 = C + 2D$. b) La mejor recta, $-2 + t$ pasa por todos los cuatro puntos; $E^2 = 0$. c) b está en el espacio columna.

3. Proyección sobre a_3: $(-2/3, 1/3, -2/3)$; la suma es b mismo; observe que $a_1a_1^{\mathrm{T}}$, $a_2a_2^{\mathrm{T}}$, $a_3a_3^{\mathrm{T}}$ son proyecciones sobre tres direcciones ortogonales. Su suma es la proyección sobre todo el espacio y debe ser la identidad.

5. $(I-2uu^{\mathrm{T}})^{\mathrm{T}}(I-2uu^{\mathrm{T}}) = I - 4uu^{\mathrm{T}} + 4uu^{\mathrm{T}}uu^{\mathrm{T}} = I$; $Q = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

7. $(x_1q_1 + \cdots + x_nq_n)^{\mathrm{T}}(x_1q_1 + \cdots + x_nq_n) = x_1^2 + \cdots + x_n^2 \Rightarrow \|b\|^2 = b^{\mathrm{T}}b = x_1^2 + \cdots + x_n^2$.

9. La combinación más próxima de q_3 es $0q_1 + 0q_2$.

11. Q es triangular superior; la columna 1 tiene $q_{11} = \pm 1$; por ortogonalidad, la columna 2 debe ser $(0, \pm 1, 0, \ldots)$; por ortogonalidad la columa 3 es $(0, 0, \pm 1, \ldots)$; y así sucesivamente.

13. $A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = QR$.

15. $q_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ -2/3 \end{bmatrix}, q_2 = \begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix}, q_3 = \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$ está en el espacio nulo izquierdo;

$\widehat{x} = \begin{bmatrix} q_1^{\mathrm{T}}b \\ q_2^{\mathrm{T}}b \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

17. $R\widehat{x} = Q^{\mathrm{T}}b$ proporciona $\begin{bmatrix} 3 & 3 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} \widehat{x} \end{bmatrix} = \begin{bmatrix} 5/3 \\ 0 \end{bmatrix}$ y $\widehat{x} = \begin{bmatrix} 5/9 \\ 0 \end{bmatrix}$.

19. $C^* - (q_2^{\mathrm{T}}C^*)\, q_2$ es $c - (q_1^{\mathrm{T}}c)\, q_1 - (q_2^{\mathrm{T}}c)\, q_2$ porque $q_2^{\mathrm{T}}q_1 = 0$.

21. Por ortogonalidad, las funciones más próximas son 0 sen $2x = 0$ y $0 + 0x = 0$.

23. $a_0 = 1/2, a_1 = 0, b_1 = 2/\pi$.

25. La recta más próxima es $y = 1/3$ (horizontal, ya que $(x, x^2) = 0$).

27. $(1/\sqrt{2}, -1/\sqrt{2}, 0, 0)$, $(1/\sqrt{6}, 1/\sqrt{6}, 2/\sqrt{6}, 0)$, $(-1/2\sqrt{3}, -1/2\sqrt{3}, 1/2\sqrt{3}, -1/\sqrt{3})$.

29. $A = a = (1, -1, 0, 0)$; $B = b - p = \left(\frac{1}{2}, \frac{1}{2}, -1, 0\right)$; $C = c - p_A - p_B = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, -1\right)$. Observe el patrón en estos vectores ortogonales A, B y C. Luego, $(1, 1, 1, 1)/4$.

31. a) Verdadero b) Verdadero. $Qx = x_1q_1 + x_2q_2$. $\|Qx\|^2 = x_1^2 + x_2^2$ porque $q_1^{\mathrm{T}}q_2 = 0$.

Conjunto de problemas 3.5, página 196

1. $F^2 = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 4 & 0 \\ 0 & 4 & 0 & 0 \end{bmatrix}$, $F^4 = \begin{bmatrix} 16 & 0 & 0 & 0 \\ 0 & 16 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 16 \end{bmatrix} = 4^2 I$.

3. La submatriz es F_3.

5. $e^{ix} = -1$ para $x = (2k+1)\pi$, $e^{i\theta} = i$ para $\theta = 2k\pi + \pi/2$, k es entero.

7. $c = (1, 0, 1, 0)$.

9. a) $y = F$ veces $(1, 0, 0, 0)$ = columna cero de $F = (1, 1, 1, 1)$.
b) $c = (1, 1, 1, 1)/4$.

11. $c = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \rightarrow \begin{matrix} c_{\text{par}} = \\ c_{\text{impar}} = \end{matrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \rightarrow \begin{matrix} y' = \\ y'' = \end{matrix} \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix} \rightarrow y = \begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \end{bmatrix}$.

13. $c_0 = (f_0 + f_1 + f_2 + f_3)/4$, $c_1 = (f_0 - if_1 - f_2 + if_3)/4$, $c_2 = (f_0 - f_1 + f_2 - f_3)/4$, $c_3 = (f_0 + if_1 - f_2 - if_3)/4$; f impar significa $f_0 = 0$, $f_2 = 0$, $f_3 = -f_1$. Luego $c_0 = 0$, $c_2 = 0$, $c_3 = -c_1$ de modo que c también es impar.

15. $F^{-1} = \begin{bmatrix} 1 & & & \\ & & 1 & \\ & 1 & & \\ & & & 1 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & 1 & & \\ 1 & i^2 & & \\ & & 1 & 1 \\ & & 1 & i^2 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & & 1 & \\ & 1 & & 1 \\ 1 & & -1 & \\ & -i & & i \end{bmatrix} = \frac{1}{4}F^{\mathrm{H}}$.

17. $D = \begin{bmatrix} 1 & & \\ & e^{2\pi i/6} & \\ & & e^{4\pi i/6} \end{bmatrix}$ y $F_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & e^{2\pi i/3} & e^{4\pi i/3} \\ 1 & e^{4\pi i/3} & e^{2\pi i/3} \end{bmatrix}$.

19. $\Lambda = \operatorname{diag}(1, i, i^2, i^3)$; $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ y P^{T} lleva a $\lambda^3 - 1 = 0$.

21. Valores característicos $e_0 = 2 - 1 - 1 = 0$, $e_1 = 2 - i - i^3 = 2$, $e_2 = 2 - (-1) - (-1) = 4$, $e_3 = 2 - i^3 - i^9 = 2$ Compruebe traza $0 + 2 + 4 + 2 = 8$.

23. Las cuatro componentes son $(c_0 + c_2) + (c_1 + c_3)$; luego, $(c_0 - c_2) + i(c_1 - c_3)$; luego, $(c_0 + c_2) - (c_1 + c_3)$; luego, $(c_0 - c_2) - i(c_1 - c_3)$. ¡Estos pasos son la **TFR**!

Conjunto de problemas 4.2, página 206

1. $\det(2A) = 8$ y $\det(-A) = (-1)^4 \det A = \frac{1}{2}$ y $\det(A^2) = \frac{1}{4}$ y $\det(A^{-1}) = 2$.

3. Por la regla 5, las operaciones en los renglones dejan sin modificar a $\det A$. Luego, al multiplicar un renglón por -1 (regla 3) se obtiene la regla del intercambio de renglones: $\det B = -\det A$.

5. Para la primera matriz, con dos intercambios de renglones se obtiene la matriz identidad. La segunda matriz requiere tres intercambios de renglones para llegar a I.

7. $\det A = 0$ (singular); $\det U = 16$; $\det U^{\mathrm{T}} = 16$; $\det U^{-1} = 1/16$; $\det M = 16$ (dos intercambios).

9. El nuevo determinante es $(1 - m\ell)(ad - bc)$.

11. Si $|\det Q|$ no es 1, entonces $\det Q^n = (\det Q)^n$ se inflaría o tendería a cero. Pero Q^n sigue siendo una matriz ortogonal. Así, $\det Q$ debe ser 1 o -1.

13. a) Con la regla 3 (factorizar -1 de cada renglón) se obtiene $\det(K^{\mathrm{T}}) = (-1)^3 \det K$. Así, con $-\det K = \det K^{\mathrm{T}} = \det K$ se obtiene $\det K = 0$.

b) $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}$ tiene det $= 1$.

15. Sumar cada columna de A a la primera columna la hace una columna cero, de modo que $\det A = 0$. Si la suma de todo renglón de A es 1, entonces la suma de todo renglón de $A - I$ es cero y $\det(A - I) = 0$.

Sin embargo, $\det A$ no debe ser 1: $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ tiene $\det(A - I) = 0$, aunque $\det A = 0 \neq 1$.

17. $\det(A) = 10$, $\det(A^{-1}) = \frac{1}{10}$, $\det(A - \lambda I) = \lambda^2 - 7\lambda + 10 = 0$ para $\lambda = 5$ y $\lambda = 2$.

19. Al tomar determinantes se obtiene $(\det C)(\det D) = (-1)^n(\det D)(\det C)$. Para n par el razonamiento fracasa (porque $(-1)^n = +1$) y la conclusión es errónea.

21. $\det(A^{-1}) = \det \begin{bmatrix} \dfrac{d}{ad-bc} & \dfrac{-b}{ad-bc} \\ \dfrac{-c}{ad-bc} & \dfrac{a}{ad-bc} \end{bmatrix} = \dfrac{ad-bc}{(ad-bc)^2} = \dfrac{1}{ad-bc}$.

23. Determinante $= 36$ y determinante $= 5$.

25. $\det(L) = 1$, $\det(U) = -6$, $\det(A) = -6$, $\det(U^{-1}L^{-1}) = -\frac{1}{6}$, y $\det(U^{-1}L^{-1}A) = 1$.

27. Renglón 3 − renglón 2 = renglón 2 − renglón 1, de modo que A es singular.

29. A es rectangular, de modo que $\det(A^{\mathrm{T}}A) \neq (\det A^{\mathrm{T}})(\det A)$: éstas no están definidas.

31. Los determinantes de Hilbert son $1, 8\times10^{-2}, 4.6\times10^{-4}, 1.6\times10^{-7}, 3.7\times10^{-12}, 5.4\times10^{-18}, 4.8\times10^{-25}, 2.7\times10^{-33}, 9.7\times10^{-43}, 2.2\times10^{-53}$. Los pivotes son razones de determinantes, así que el décimo pivote está próximo a $10^{-53}/10^{-43} = 10^{-10}$: demasiado pequeño.

33. Los determinantes más grandes de las matrices 0−1 para $n = 1, 2, \ldots,$ son 1, 1, 2, 3, 5, 9, 32, 56, 144, 320, en la página web *www.mathworld.wolfram.com/Hadamards-MaximumDeterminantProblem.htm/* y también en la "On−Line Encyclopedia of Integer Sequences"; *www.research.att.com*. Con −1s y 1s, el determinante de 4 por 4 más grande es 16 (consulte Hadamard en el índice).

35. $\det(I + M) = 1 + a + b + c + d$. Reste el renglón 4 de los renglones 1, 2 y 3. Luego, reste a(renglón 1) + b(renglón 2) + c(renglón 3) del renglón 4. Al hacer lo anterior, se queda con una matriz triangular con 1, 1, 1, y $1 + a + b + c + d$ en su diagonal.

Conjunto de problemas 4.3, página 215

1. a) $a_{12}a_{21}a_{34}a_{43} = 1$; *par*, de modo que $\det A = 1$.
b) $b_{13}b_{22}b_{31}b_{14} = 18$; *impar*, de modo que $\det B = -18$.

3. a) Verdadero (regla del producto). b) Falso (todo 1s).
c) Falso ($\det[1\ \ 1\ \ 0;\ 0\ \ 1\ \ 1;\ 1\ \ 0\ \ 1] = 2$).

5. El cofactor 1, 1, es F_{n-1}. El cofactor 1, 2 tiene un 1 en la columna 1, con cofactor F_{n-2}. Multiplique por $(-1)^{1+2}$ y también −1 para encontrar $F_n = F_{n-1} + F_{n-2}$. Así, los determinantes son números de Fibonacci, excepto que F_n es el F_{n-1} de costumbre.

7. Desarrollo por cofactores: $\det = 4(3) - 4(1) + 4(-4) - 4(1) = -12$.

9. a) $(n-1)n!$ (cada término $n - 1$). b) $\left(1 + \dfrac{1}{2!} + \cdots + \dfrac{1}{(n-1)!}\right)n!$
c) $\dfrac{1}{3}(n^3 + 2n - 3)$.

11. $\begin{bmatrix} 0 & A \\ -B & I \end{bmatrix}\begin{bmatrix} I & 0 \\ B & I \end{bmatrix} = \begin{bmatrix} AB & A \\ 0 & I \end{bmatrix}$, $\det\begin{bmatrix} I & 0 \\ B & I \end{bmatrix} = 1 \Rightarrow \det\begin{bmatrix} 0 & A \\ -B & I \end{bmatrix} = \det\begin{bmatrix} AB & A \\ 0 & I \end{bmatrix} = \det(AB)$. Compruebe $A = [1\ \ 2]$, $B = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\det\begin{bmatrix} 0 & A \\ -B & I \end{bmatrix} = 5 = \det(AB)$; $A = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $B = [1\ \ 2]$, $\det\begin{bmatrix} 0 & A \\ -B & I \end{bmatrix} = 0 = \det(AB)$. Singular: rango $(AB) \leq$ rango $(A) \leq n < m$.

13. $\det A = 1 + 18 + 12 - 9 - 4 - 6 = 12$, de modo que los renglones son independientes; $\det B = 0$, por lo que los renglones son dependientes (renglón 1 + renglón 2 = renglón 3); $\det C = -1$, C tiene renglones independientes.

15. Cada uno de los seis términos en $\det A$ es cero; el rango es cuando mucho 2; la columna 2 carece de pivote.

17. $a_{11}a_{23}a_{32}a_{44}$ tiene −, $a_{14}a_{23}a_{32}a_{41}$ tiene +, de modo que $\det A = 0$; $\det B = 2 \cdot 4 \cdot 4 \cdot 2 - 1 \cdot 4 \cdot 4 \cdot 1 = 48$.

19. a) Si $a_{11} = a_{22} = a_{33} = 0$, entonces es seguro que cuatro términos son cero.
b) Quince términos son cero.

21. ¡Algún término $a_{1\alpha}a_{2\beta}\cdots a_{n\omega}$ en la gran fórmula es diferente de cero! Mueva los renglones $1, 2, \ldots, n$ hacia los renglones $\beta, \ldots, \omega$. Por tanto, estas as distintas de cero están sobre la diagonal principal.

23. $4!/2 = 12$ permutaciones pares; $\det(I + P_{\text{par}}) = 16$ o 4 o 0 (16 proviene de $I + I$).

25. $C = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}$ y $AC^{\mathrm{T}} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} = 4I$. En consecuencia, $A^{-1} = \dfrac{1}{4}C^{\mathrm{T}}$.

27. $|B_n| = |A_n| - |A_{n-1}| = (n+1) - n = 1$.

29. Es necesario escoger 1s de las columnas 2 y 1, de las columnas 4 y 3 y así sucesivamente. En consecuencia, n debe ser par a fin de tener $\det A_n \neq 0$. El número de intercambios es $\frac{1}{2}n$, por lo que $C_n = (-1)^{n/2}$.

31. $S_1 = 3$, $S_2 = 8$, $S_3 = 21$. Parece que la regla es que cada segundo número en la sucesión de Fibonacci . . . , 3, 5, 8, 13, 21, 34, 55, . . . , por lo que la conjetura es $S_4 = 55$. Los cinco términos diferentes de cero en la gran fórmula para S_4 son (con números 3 donde el problema 39 tiene números 2) $81 + 1 - 9 - 9 - 9 = 55$.

33. Al cambiar 3 por 2 en el vértice el determinante F_{2n+2} se reduce por 1 multiplicado por el cofactor de ese elemento del vértice. Este cofactor es el determinante de S_{n-1} (un tamaño menor), que es F_{2n}. En consecuencia, al cambiar 3 por 2 se modifica el determinante a $F_{2n+2} - F_{2n}$, que es F_{2n+1}.

35. a) Todo $\det L = 1$; $\det U_k = \det A_k = 2, 6, -6$ para $k = 1, 2, 3$.
b) Pivotes $5, \frac{6}{5}, \frac{7}{6}$.

37. Los seis términos son correctos. Renglón 1 − 2 renglón 2 + renglón 3 = 0, de modo que la matriz es singular.

39. Los cinco términos diferentes de cero en $\det A = 5$ son
$(2)(2)(2)(2) + (-1)(-1)(-1)(-1) - (-1)(-1)(2)(2) - (2)(2)(-1)(-1) - (2)(-1)(-1)(2)$.

41. Con $a_{11} = 1$, el determinante de la matriz $-1, 2, -1$ es det $= 1$ y la inversa $(A^{-1})_{ij} = n + 1 - \text{máx}\,(i, j)$.

43. Al restar 1 del elemento n, n se resta su cofactor C_{nn} del determinante. Este cofactor es $C_{nn} = 1$ (la menor matriz de Pascal). Al restar 1 de 1 se obtiene 0.

Conjunto de problemas 4.4, página 225

1. $\det A = 20; C^{\mathrm{T}} = \begin{bmatrix} 20 & -10 & -12 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{bmatrix}; AC^{\mathrm{T}} = 20I; A^{-1} = \dfrac{1}{20}\begin{bmatrix} 20 & -10 & -12 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{bmatrix}$.

3. $(x, y) = (d/(ad - bc), -c/(ad - bc)); (x, y, z) = (3, -1, -2)$.

5. a) El área de ese paralelogramo es $\det \begin{bmatrix} 2 & 2 \\ -1 & 3 \end{bmatrix}$, por lo que el área del triángulo ABC es $\frac{1}{2}4 = 2$. b) El triángulo $A'B'C'$ tiene la misma área; simplemente se ha movido hacia el origen.

7. Los pivotes de A son 2, 3 y 6 a partir de los determinantes 2, 3, y 36; los pivotes de B son 2, 3 y 0.

9. a) P^2 lleva (1, 2, 3, 4, 5) a (3, 2, 5, 4, 1).
b) P^{-1} lleva (1, 2, 3, 4, 5) a (3, 4, 5, 2, 1).

11. Las potencias de P son todas las matrices permutación, por lo que finalmente una de esas matrices debe repetirse. Si P^r es la misma que P^s, entonces $P^{r-s} = I$.

13. a) $\det A = 3$, $\det B_1 = -6$, $\det B_2 = 3$, de modo que $x_1 = -6/3 = -2$ y $x_2 = 3/3 = 1$. b) $|A| = 4$, $|B_1| = 3$, $|B_2| = -2$, $|B_3| = 1$. Por tanto, $x_1 = \frac{3}{4}$, $x_2 = -\frac{1}{2}$ y $x_3 = \frac{1}{4}$.

15. a) $x_1 = \frac{3}{0}$ y $x_2 = \frac{-2}{0}$: no hay solución. b) $x_1 = \frac{0}{0}$ y $x_2 = \frac{0}{0}$: *indeterminada*.

17. Si la primera columna de A es también el miembro derecho b, entonces $\det A = \det B_1$. Tanto B_2 como B_3 son singulares, ya que se repite una columna. En consecuencia, $x_1 = |B_1|/|A| = 1$ y $x_2 = x_3 = 0$.

19. Si todos los cofactores $= 0$ (incluso en un solo renglón o en una sola columna), entonces $\det A = 0$ (no hay inversa). $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ no tiene cofactores cero pero no es invertible.

21. Si $\det A = 1$ y se conocen los cofactores, entonces $C^{\mathrm{T}} = A^{-1}$ y también $\det A^{-1} = 1$. Debido a que A es la inversa de A^{-1}, A debe ser la matriz de cofactores de C.

23. Una vez que se conoce C, el problema 22 proporciona $\det A = (\det C)^{\frac{1}{n-1}}$ con $n = 4$. Así, es posible construir $A^{-1} = C^{\mathrm{T}}/\det A$ usando los cofactores conocidos. Para encontrar A es necesario invertir.

25. a) Cofactores $C_{21} = C_{31} = C_{32} = 0$.
b) $C_{12} = C_{21}$, $C_{31} = C_{13}$, $C_{32} = C_{23}$ hace simétrica a S^{-1}.

27. a) Área $\begin{vmatrix} 3 & 2 \\ 1 & 4 \end{vmatrix} = 10$. b) Área del triángulo $= 5$. c) Área del triángulo $= 5$.

29. a) Área $\frac{1}{2}\begin{vmatrix} 2 & 1 & 1 \\ 3 & 4 & 1 \\ 0 & 5 & 1 \end{vmatrix} = 5$. b) $5 +$ nueva área del triángulo $\frac{1}{2}\begin{vmatrix} 2 & 1 & 1 \\ 0 & 5 & 1 \\ -1 & 0 & 1 \end{vmatrix} = 5 + 7 = 12$.

31. Las aristas del hipercubo tienen longitud $\sqrt{1+1+1+1} = 2$. El volumen $\det H$ es $2^4 = 16$. ($H/2$ tiene columnas ortonormales. Así, $\det(H/2) = 1$ lleva de nuevo a $\det H = 16$).

33. $x_1y_2 - x_2y_1 =$ rectángulos $A + B + D$ (not C).

Areas de mismas bases y mismas alturas: rectangulo $A = 2$(triángulo a), rectangulo $B = 2$(triángulo b), rectangulo $D = 2$(triángulo d).

Por lo que los triangulos $a + b + d = \frac{1}{2}(x_1y_2 - x_2y_1)$.

Compruebe un ejemplo con $(a, b) = (3, 2)$, $(c, d) = (1, 4)$ y área $= 10$. La recta desde $(0, e)$ en el paso 3 tiene pendiente c/a y su ecuación es $y = e + cx/a$. ¡El paso 3 funciona porque (b, d) está en esa recta! $d = e + cb/a$ es verdadera, ya que $ad - bc =$ área ae en el paso 2.

35. El cubo n-dimensional tiene 2^n vértices, $n2^{n-1}$ aristas y $2n$ caras de dimensión $n - 1$. El volumen del cubo cuyas aristas son los renglones de $2I$ es 2^n.

37. $J = r$. Las columnas son ortogonales y sus longitudes son 1 y r.

39. $\begin{vmatrix} \partial r/\partial x & \partial r/\partial y \\ \partial\theta/\partial x & \partial\theta/\partial y \end{vmatrix} = \begin{vmatrix} \cos\theta & \operatorname{sen}\theta \\ (-\operatorname{sen}\theta)/r & (\cos\theta)/r \end{vmatrix} = \dfrac{1}{r}$.

41. $S = (2, 1, -1)$ proporciona un paralelogramo, cuya área es la longitud de un producto cruz: $\|PQ \times PS\| = \|(-2, -2, -1)\| = 3$. ¡Esto también proviene de un determinante! Los otros cuatro vértices podrían ser (0, 0, 0), (0, 0, 2), (1, 2, 2) y (1, 1, 0). El volumen de la caja inclinada es $|\det| = 1$.

43. $\det \begin{bmatrix} x & y & z \\ 3 & 2 & 1 \\ 1 & 2 & 3 \end{bmatrix} = 0 = 7x - 5y + z$; el plano contiene dos vectores.

45. VISA tiene cinco inversiones VI, VS, VA, IA, SA. Y AVIS tiene dos inversiones VI y VS. Debido a que $5 - 2$ es impar, VISA y AVIS tiene paridad opuesta.

Conjunto de problemas 5.1, página 240

1. $\lambda = 2$ y $\lambda = 3$; traza $= 5$, determinante $= 6$.

3. $\lambda = -5$ y $\lambda = -4$; las dos λs son reducidas por 7, con vectores característicos sin modificar.

5. $\lambda = 3, \lambda = 1, \lambda = 0$, con vectores característicos (1, 0, 0), (2, −1, 0), (3, −2, 1); traza $= 4$, det $= 0$. $\lambda = 2, \lambda = 2, \lambda = -2$, con vectores característicos (1, 1, 1), (0, 1, 0), (1, 0, −1); traza $= 2$, det $= -8$.

7. $Ax = \lambda x$ proporciona $(A - 7I)x = (\lambda - 7)x$; $Ax = \lambda x$ proporciona $x = \lambda A^{-1}x$, de modo que $A^{-1}x = (1/\lambda)x$.

9. El coeficiente de $(-\lambda)^{n-1}$ en $(\lambda_1 - \lambda)\cdots(\lambda_n - \lambda)$ es $\lambda_1 + \cdots + \lambda_n$. En $\det(A - \lambda I)$, un término que incluye un a_{ij} fuera de la diagonal excluye tanto a $a_{ii} - \lambda$ como a $a_{jj} - \lambda$. Este término no implica a $(-\lambda)^{n-1}$. Así, el coeficiente de $(-\lambda)^{n-1}$ en $\det(A - \lambda I)$ debe provenir del producto bajo la diagonal principal. Ese coeficiente es $a_{11} + \cdots + a_{nn} = \lambda_1 + \cdots + \lambda_n$.

11. Trasponga $A - \lambda I$: $\det(A - \lambda I) = \det(A - \lambda I)^{\mathrm{T}} = \det(A^{\mathrm{T}} - \lambda I)$.

13. Los valores característicos de A son 1, 2, 3, 7, 8, 9.

15. rango$(A) = 1$, $\lambda = 0, \ldots, 0, n$ (traza n); rango$(C) = 2$, $\lambda = 0, \ldots, n/2, -n/2$ (traza 0).

17. El tercer renglón contiene a 6, 5, 4.

19. A, A^2 y A^∞ tienen los mismos vectores característicos. Los valores característicos son 1 y 0.5 para A, 1 y 0.25 para A^2 y 1 y 0 para A^∞. En consecuencia, A^2 está a medio camino entre A y A^∞.

21. $\lambda_1 = 4$ y $\lambda_2 = -1$ (compruebe la traza y el determinante) con $x_1 = (1, 2)$ y $x_2 = (2, -1)$. A^{-1} tiene los mismos vectores característicos que A, con valores característicos $1/\lambda_1 = 1/4$ y $1/\lambda_2 = -1$.

23. a) Multiplique Ax para ver λx, lo cual revela a λ. b) Resuelva $(A - \lambda I)x = 0$ para encontrar x.

25. a) $Pu = (uu^T)u = u(u^Tu) = u$, de modo que $\lambda = 1$. b) $Pv = (uu^T)v = u(u^Tv) = 0$, así que $\lambda = 0$. c) $x_1 = (-1, 1, 0, 0)$, $x_2 = (-3, 0, 1, 0)$, $x_3 = (-5, 0, 0, 1)$ son ortogonales a u, por lo que son vectores característicos de P con $\lambda = 0$.

27. $\lambda^3 - 1 = 0$ proporciona $\lambda = 1$ y $\lambda = \frac{1}{2}(-1 \pm i\sqrt{3})$; los tres valores característicos son $1, 1, -1$.

29. a) rango = 2. b) $\det(B^TB) = 0$. *No* c). d) $(B + I)^{-1}$ tiene $(\lambda + 1)^{-1} = 1, \frac{1}{2}, \frac{1}{3}$.

31. $a = 0$, $b = 9$, $c = 0$ multiplican a $1, \lambda, \lambda^2$ en $\det(A - \lambda I) = 9\lambda - \lambda^3$: A = *matriz acompañante*.

33. $\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}$, $\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}$, $\begin{bmatrix}-1 & 1\\-1 & 1\end{bmatrix}$. Siempre A^2 = matriz cero si $\lambda = 0, 0$ (Cayley-Hamilton).

35. $Ax = c_1\lambda_1x_1 + \cdots + c_n\lambda_nx_n$ es igual a $Bx = c_1\lambda_1x_1 + \cdots + c_n\lambda_nx_n$ para toda x. Así, $A = B$.

37. $\begin{bmatrix}a & b\\c & d\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}a+b\\c+d\end{bmatrix} = (a+b)\begin{bmatrix}1\\1\end{bmatrix}$; $\lambda_2 = d - b$ para obtener traza $= a + d$.

39. Se requiere $\lambda^3 = 1$, pero no $\lambda = 1$ (para evitar I). Con $\lambda_1 = e^{2\pi i/3}$ y $\lambda_2 = e^{-2\pi i/3}$, el determinante es $\lambda_1\lambda_2 = 1$ y la traza es $\lambda_1 + \lambda_2 = \cos\frac{2\pi}{3} + i\,\text{sen}\frac{2\pi}{3} + \cos\frac{2\pi}{3} - i\,\text{sen}\frac{2\pi}{3} = -1$. Una matriz con esta traza -1 y determinante 1 es $A = \begin{bmatrix}-1 & 1\\-1 & 0\end{bmatrix}$.

Conjunto de problemas 5.2, página 250

1. $\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix} = \begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}\begin{bmatrix}2 & 0\\0 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}^{-1}$;

$\begin{bmatrix}2 & 1\\0 & 0\end{bmatrix} = \begin{bmatrix}1 & 1\\0 & -2\end{bmatrix}\begin{bmatrix}2 & 0\\0 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\0 & -2\end{bmatrix}^{-1}$.

3. $\lambda = 0, 0, 3$; la tercera columna de S es un múltiplo de $\begin{bmatrix}1\\1\\1\end{bmatrix}$ y las otras columnas están sobre el plano ortogonal a aquélla.

5. No es posible diagonalizar a A_1 y A_3. Sólo tienen una recta de vectores característicos.

7. $A = \begin{bmatrix}3 & 1\\1 & -1\end{bmatrix}\begin{bmatrix}5 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}3 & 1\\1 & -1\end{bmatrix}^{-1}$ proporciona

$A^{100} = \begin{bmatrix}3 & 1\\1 & -1\end{bmatrix}\begin{bmatrix}5^{100} & 0\\0 & 1\end{bmatrix}\begin{bmatrix}3 & 1\\1 & -1\end{bmatrix}^{-1} = \frac{1}{4}\begin{bmatrix}3\cdot 5^{100} + 1 & 3\cdot 5^{100} - 3\\5^{100} - 1 & 5^{100} + 3\end{bmatrix}$.

9. Traza(AB) = traza $(BA) = aq + bs + cr + dt$. Así, traza $(AB - BA) = 0$ (siempre). Por tanto, $AB - BA = I$ es imposible para matrices, ya que la traza de I no es cero.

11. a) Verdadero; $\det A = 2 \neq 0$. b) Falso; $\begin{bmatrix}1 & 1 & 1\\0 & 1 & 1\\0 & 0 & 2\end{bmatrix}$. c) Falso; $\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 2\end{bmatrix}$ es diagonal!

13. $A = \begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}\begin{bmatrix}9 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}^{-1}$; $\begin{bmatrix}2 & 1\\1 & 2\end{bmatrix}$; cuatro raíces cuadradas.

15. $\begin{bmatrix}1 & 2\\0 & 3\end{bmatrix} = \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 3\end{bmatrix}\begin{bmatrix}1 & -1\\0 & 1\end{bmatrix}$; $\begin{bmatrix}1 & 1\\2 & 2\end{bmatrix} = \begin{bmatrix}1 & 1\\-1 & 2\end{bmatrix}\begin{bmatrix}0 & 0\\0 & 3\end{bmatrix}\begin{bmatrix}\frac{2}{3} & -\frac{1}{3}\\\frac{1}{3} & \frac{1}{3}\end{bmatrix}$.

17. $A = \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}\begin{bmatrix}2 & 0\\0 & 5\end{bmatrix}\begin{bmatrix}1 & -1\\0 & 1\end{bmatrix} = \begin{bmatrix}2 & 3\\0 & 5\end{bmatrix}$.

19. a) Falso; no se conocen λs. b) Verdadero. c) Verdadero. d) Falso: ¡se requieren vectores característicos de S!

21. Las columnas de S son múltiplos de (2, 1) y (0, 1) en cualquier orden. Lo mismo para A^{-1}.

23. A y B tienen $\lambda_1 = 1$ y $\lambda_2 = 1$. $A + B$ tiene $\lambda_1 = 1$ y $\lambda_2 = 3$. Los valores característicos de $A + B$ *no son iguales* a los valores característicos de A más los valores característicos de B.

25. a) Verdadero. b) Falso. c) Falso (A podría tener 2 o 3 vectores característicos independientes).

27. $A = \begin{bmatrix}8 & 3\\-3 & 2\end{bmatrix}$ (u otra), $A = \begin{bmatrix}9 & 4\\-4 & 1\end{bmatrix}$, $A = \begin{bmatrix}10 & 5\\-5 & 0\end{bmatrix}$; los únicos vectores característicos son $(c, -c)$.

29. $S\Lambda^kS^{-1}$ tiende a cero si y sólo si todo $|\lambda| < 1$; $B^k \to 0$ desde $\lambda = 0.9$ y $\lambda = 0.3$.

31. $\Lambda = \begin{bmatrix}0.9 & 0\\0 & 0.3\end{bmatrix}$, $S = \begin{bmatrix}3 & -3\\1 & 1\end{bmatrix}$; $B^{10}\begin{bmatrix}3\\1\end{bmatrix} = (0.9)^{10}\begin{bmatrix}3\\1\end{bmatrix}$, $B^{10}\begin{bmatrix}3\\-1\end{bmatrix} = (0.3)^{10}\begin{bmatrix}3\\-1\end{bmatrix}$,

$B^{10}\begin{bmatrix}6\\0\end{bmatrix}$ = suma de estas dos.

33. $B^k = \begin{bmatrix}1 & 1\\0 & -1\end{bmatrix}\begin{bmatrix}3 & 0\\0 & 2\end{bmatrix}^k\begin{bmatrix}1 & 1\\0 & -1\end{bmatrix} = \begin{bmatrix}3^k & 3^k - 2^k\\0 & 2^k\end{bmatrix}$.

35. Traza $AB = (aq + bs) + (cr + dt) = (qa + rc) + (sb + td)$ = traza BA. Demostración para el caso diagonalizable: la traza de $S\Lambda S^{-1}$ es la traza de $(\Lambda S^{-1})S = \Lambda$, que es *la suma de los* λ*s*.

37. Las As forman un subespacio, ya que cA y $A_1 + A_2$ tienen la misma S. Cuando $S = I$, las As proporcionan el subespacio de las matrices diagonales. Dimensión 4.

39. Dos problemas: el espacio nulo y el espacio columna pueden traslaparse, por lo que x podría estar en ambos. En el espacio columna podría no haber r vectores característicos independientes.

41. $A = \begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}$ tiene $A^2 = \begin{bmatrix}2 & 1\\1 & 1\end{bmatrix}$, y $A^2 - A - I$ = matriz cero confirma el teorema de Cayley-Hamilton.

43. Por 5F, B tiene los mismos vectores característicos (1, 0) y (0, 1) que A, por lo que B es diagonal. Las ecuaciones $AB - BA = \begin{bmatrix} a & b \\ 2c & 2d \end{bmatrix} - \begin{bmatrix} a & 2b \\ c & 2d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ son $-b = 0$ y $c = 0$: rango 2.

45. A tiene $\lambda_1 = 1$ y $\lambda_2 = 0.4$ con $x_1 = (1, 2)$ y $x_2 = (1, -1)$. A^∞ tiene $\lambda_1 = 1$ y $\lambda_2 = 0$ (mismos vectores característicos). A^{100} tiene $\lambda_1 = 1$ y $\lambda_2 = (0.4)^{100}$, que está próximo a cero.

Conjunto de problemas 5.3, página 262

1. Los números de Fibonacci comienzan par, impar, impar. Así, *impar* + *impar* = *par*. Los dos siguientes son impar (de impar + par y par + impar). Luego se repite *impar* + *impar* = *par*.

3. $A^2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, $A^3 = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}$, $A^4 = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$; $F_{20} = 6765$.

5. $A = S\Lambda S^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \dfrac{1}{\lambda_1 - \lambda_2} \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{bmatrix}$ (observe S^{-1}).

$$S\Lambda^k S^{-1} = \frac{1}{\lambda_1 - \lambda_2} \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{bmatrix} \begin{bmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} ------ \\ (\lambda_1^k - \lambda_2^k)/(\lambda_1 - \lambda_2) \end{bmatrix}.$$

7. La suma directa $L_k + L_{k+1}$ proporciona $L_0, \ldots, L_{10}$ como 2, 1, 3, 4, 7, 11, 18, 29, 47, 76, 123. Con mi calculadora obtengo $\lambda_1^{10} = (1.618\ldots)^{10} = 122.991\ldots$, que se redondea a $L_{10} = 123$.

9. La matriz de transición de Markov es $\begin{bmatrix} \frac{7}{12} & \frac{1}{6} & 0 \\ \frac{1}{6} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{3} & 1 \end{bmatrix}$. Las fracciones $\frac{7}{12}$, $\frac{1}{2}$ y 1 no se mueven.

11. a) $\lambda = 0$, $(1, 1, -2)$. b) $\lambda = 1$ y -0.2. c) límite (3, 4, 4) = vector característico para $\lambda = 1$.

13. a) $0 \le a \le 1$, $0 \le b \le 1$.

b) $$u_k = \begin{bmatrix} b/(1-a) & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1^k & 0 \\ 0 & (a-b)^k \end{bmatrix} \begin{bmatrix} b/(1-a) & 1 \\ 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \dfrac{2b}{b-a+1} - \dfrac{1-a-b}{b-a+1}(a-b)^k \\ \dfrac{2(1-a)}{b-a+1} - \dfrac{1-a-b}{b-a+1}(a-b)^k \end{bmatrix}.$$

c) $u_k \to \begin{bmatrix} \dfrac{2b}{b-a+1} \\ \dfrac{2(1-a)}{b-a+1} \end{bmatrix}$ si $|a - b| < 1$; $a = 1/3$, $b = -1/3$ no es de Markov.

15. La suma de las componentes de Ax es $x_1 + x_2 + x_3$ (la suma de cada columna es 1 y no se pierde nada). La suma de las componentes de λx es $\lambda(x_1 + x_2 + x_3)$. Si $\lambda \neq 1$, $x_1 + x_2 + x_3$ debe ser cero.

17. $\begin{bmatrix} \alpha & \alpha \\ \alpha & \alpha \end{bmatrix}$ es inestable para $|\alpha| > 1/2$, y estable para $|\alpha| < 1/2$. Es neutra para $\alpha = \pm 1/2$.

19. $A^2 = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ y $A^3 = 0$. Así, $(I - A)^{-1} = I + A + A^2 = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

21. Si A se incrementa, entonces se consumen más bienes en la producción y la expansión debe ser más lenta. En términos matemáticos, $Ax \ge tx$ debe seguir siendo verdadera si A se incrementa; $t_{máx}$ se va para arriba.

23. $\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ y $A^k = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 5^k & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$.

25. $R = S\sqrt{\Lambda}S^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ tiene $R^2 = A$. $\sqrt{B}$ debería tener $\lambda = \sqrt{9}$ y $\lambda = \sqrt{-1}$, por lo que su traza no es real. Observe que $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ puede tener $\sqrt{-1} = i$ y $-i$, y una raíz cuadrada real $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

27. $A = S\Lambda_1 S^{-1}$ y $B = S\Lambda_2 S^{-1}$. Las matrices diagonales siempre proporcionan $\Lambda_1\Lambda_2 = \Lambda_2\Lambda_1$. Así, $AB = BA$, a partir de $S\Lambda_1 S^{-1} S\Lambda_2 S^{-1} = S\Lambda_1\Lambda_2 S^{-1} = S\Lambda_2\Lambda_1 S^{-1} = S\Lambda_2 S^{-1} S\Lambda_1 S^{-1} = BA$.

29. B tiene $\lambda = i$ y $-i$, de modo que B^4 tiene $\lambda^4 = 1$ y 1; C tiene $\lambda = (1 \pm \sqrt{3}i)/2 = \exp(\pm \pi i/3)$, por lo que $\lambda^3 = -1$ y -1. Así, $C^3 = -I$ y $C^{1024} = -C$.

Conjunto de problemas 5.4, página 275

1. $\lambda_1 = -2$ y $\lambda_2 = 0$; $x_1 = (1, -1)$ y $x_2 = (1, 1)$; $e^{At} = \frac{1}{2} \begin{bmatrix} e^{-2t} + 1 & -e^{-2t} + 1 \\ -e^{-2t} + 1 & e^{-2t} + 1 \end{bmatrix}$.

3. $u(t) = \begin{bmatrix} e^{2t} + 2 \\ -e^{2t} + 2 \end{bmatrix}$; cuando $t \to \infty$, $e^{2t} \to +\infty$.

5. a) $e^{A(t+T)} = Se^{\Lambda(t+T)}S^{-1} = Se^{\Lambda t}e^{\Lambda T}S^{-1} = Se^{\Lambda t}S^{-1}Se^{\Lambda T}S^{-1} = e^{At}e^{AT}$.

b) $e^A = I + A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, $e^B = I + B = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$, $A + B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ proporciona $e^{A+B} = \begin{bmatrix} \cos 1 & -\mathrm{sen}\, 1 \\ \mathrm{sen}\, 1 & \cos 1 \end{bmatrix}$ a partir del ejemplo 3 en el texto, en $t = 1$. Esta matriz es diferente de $e^A e^B$.

7. $e^{At} = I + At = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$; $e^{\Lambda t}u(0) = \begin{bmatrix} 4t + 3 \\ 4 \end{bmatrix}$.

9. a) $\lambda_1 = \frac{7+\sqrt{57}}{2}$, $\lambda_2 = \frac{7-\sqrt{57}}{2}$, Re $\lambda_1 > 0$, inestable. b) $\lambda_1 = \sqrt{7}$, $\lambda_2 = -\sqrt{7}$, Re $\lambda_1 > 0$, inestable c) $\lambda_1 = \frac{-1+\sqrt{13}}{2}$, $\lambda_2 = \frac{-1-\sqrt{13}}{2}$, Re $\lambda_1 > 0$, inestable d) $\lambda_1 = 0$, $\lambda_2 = -2$, neutralmente estable.

11. A_1 es inestable para $t < 1$, neutralmente estable para $t \ge 1$. A_2 es inestable para $t < 4$, neutralmente estable en $t = 4$, estable con λ real para $4 < t \le 5$, y estable con λ complejo para $t > 5$. A_3 es inestable para toda $t > 0$, porque la traza es $2t$.

13. a) $u_1' = cu_2 - bu_3$, $u_2' = -cu_1 + au_3$, $u_3' = bu_1 - au_2$ proporciona $u_1'u_1 + u_2'u_2 + u_3'u_3 = 0$. b) Debido a que e^{At} es una matriz ortogonal, $\|u(t)\|^2 = \|e^{At}u(0)\|^2 = \|u(0)\|^2$ es

constante. c) $\lambda = 0$ y $\pm(\sqrt{a^2+b^2+c^2})i$. Las matrices simétricas sesgadas tienen λs imaginarios puros.

15. $u(t) = \frac{1}{2}\cos 2t \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \frac{1}{2}\cos\sqrt{6}t \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

17. $Ax = \lambda F x + \lambda^2 x$ o bien, $(A - \lambda F - \lambda^2 I)x = 0$.

19. Los valores característicos son reales cuando $(\text{traza})^2 - 4\det \geq 0 \Rightarrow -4(-a^2 - b^2 + c^2) \geq 0 \Rightarrow a^2 + b^2 \geq c^2$.

21. $u_1 = e^{4t}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $u_2 = e^t \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Si $u(0) = (5, -2)$, entonces $u(t) = 3e^{4t}\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2e^t \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

23. $\begin{bmatrix} y' \\ y'' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} y \\ y' \end{bmatrix}$. Así se obtiene $\lambda = \frac{1}{2}(5 \pm \sqrt{41})$.

25. $\lambda_1 = 0$ y $\lambda_2 = 2$. Luego, $v(t) = 20 + 10e^{2t} \to \infty$ cuando $t \to \infty$.

27. $A = \begin{bmatrix} 0 & 1 \\ -9 & 6 \end{bmatrix}$ tiene traza 6, det 9, $\lambda = 3$ y 3, con un solo vector característico independiente $(1, 3)$. Así se obtiene $y = ce^{3t}$, $y' = 3e^{3t}$. También te^{3t} resuelve $y'' = 6y' - 9y$.

29. $y(t) = \cos t$ empieza en $y(0) = 1$ y $y'(0) = 0$. La ecuación vectorial tiene $u = (y, y') = (\cos t, -\text{sen}\, t)$.

31. Al sustituir $u = e^{ct}v$ se obtiene $ce^{ct}v = Ae^{ct}v - e^{ct}b$, o bien, $(A - cI)v = b$, o $v = (A - cI)^{-1}b$ = solución particular. Si c es un valor característico, entonces $A - cI$ no es invertible: esta v fracasa.

33. $de^{At}/dt = A + A^2 t + \frac{1}{2}A^3t^2 + \frac{1}{6}A^4t^3 + \cdots = A(I + At + \frac{1}{2}A^2t^2 + \frac{1}{6}A^3t^3 + \cdots) = Ae^{At}$.

35. La solución en el instante $t + T$ también es $e^{A(t+T)}u(0)$. Así, e^{At} multiplicada por e^{AT} es igual a $e^{A(t+T)}$.

37. Si $A^2 = A$, entonces $e^{At} = I + At + \frac{1}{2}At^2 + \frac{1}{6}At^3 + \cdots = I + (e^t - 1)A$
$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} e^t - 1 & e^t - 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} e^t & e^t - 1 \\ 0 & 1 \end{bmatrix}$.

39. $A = \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{bmatrix}$, entonces $e^{At} = \begin{bmatrix} e^t & \frac{1}{2}(e^{3t} - e^t) \\ 0 & e^{3t} \end{bmatrix} = I$ en $t = 0$.

41. a) La inversa de e^{At} es e^{-At}. b) Si $Ax = \lambda x$, entonces $e^{At}x = e^{\lambda t}x$ y $e^{\lambda t} \neq 0$.

43. $\lambda = 2$ y 5 con vectores característicos $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Así, $A = S\Lambda S^{-1} = \begin{bmatrix} -1 & 6 \\ -3 & 8 \end{bmatrix}$.

Conjunto de problemas 5.5, página 288

1. b) suma $= 4 + 3i$; producto $= 7 + i$. c) $\overline{3 + 4i} = 3 - 4i$; $\overline{1 - i} = 1 + i$; $|3 + 4i| = 5$; $|1 - i| = \sqrt{2}$. Ambos números están *fuera* de la circunferencia unitaria.

3. $\bar{x} = 2 - i$, $x\bar{x} = 5$, $xy = -1 + 7i$, $1/x = 2/5 - (1/5)i$, $x/y = 1/2 - (1/2)i$; compruebe que $|xy| = \sqrt{50} = |x||y|$ y $|1/x| = 1/\sqrt{5} = 1/|x|$.

5. a) $x^2 = r^2 e^{i2\theta}$, $x^{-1} = (1/r)e^{-i\theta}$, $\bar{x} = re^{-i\theta}$; $x^{-1} = \bar{x}$ proporciona $|x|^2 = 1$: en la circunferencia unitaria.

7. $C = \begin{bmatrix} 1 & -i \\ -i & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & i & 0 \\ i & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & i & -i \\ -i & 1 & 0 \\ i & 0 & 1 \end{bmatrix}$, $C^H = C$ porque $(A^H A)^H = A^H A$.

9. a) $\det A^T = \det A$ pero $\det A^H = \overline{\det A}$. b) $A^H = A$ proporciona $\overline{\det A} = \det A$ = real.

11. $P: \lambda_1 = 0$, $\lambda_2 = 1$, $x_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$, $x_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$; $Q: \lambda_1 = 1$, $\lambda_2 = -1$, $x_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$, $x_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$; $R: \lambda_1 = 5$, $\lambda_2 = -5$, $x_1 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$, $x_2 = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$.

13. a) u, v y w son ortogonales entre sí. b) El espacio nulo es generado por u; el espacio nulo izquierdo es el mismo que el espacio nulo; el espacio renglón es generado por v y w; el espacio columna es el mismo que el espacio renglón. c) $x = v - \frac{1}{2}w$; no es único, ya que es posible sumar a x cualquier múltiplo de u. d) Se requiere $b^T u = 0$. e) $S^{-1} = S^T$; $S^{-1}AS = \text{diag}(0, 1, 2)$.

15. La dimensión de **S** es $n(n+1)/2$, no n. Toda matriz simétrica A es una combinación de n proyecciones, pero las proyecciones cambian cuando A cambia. En el espacio **S** de matrices simétricas no hay base de n matrices proyección fijas.

17. $(UV)^H(UV) = V^H U^H U V = V^H I V = I$. Así, UV es unitaria.

19. La tercera columna de U puede ser $(1, -2, i)/\sqrt{6}$, multiplicado por cualquier número $e^{i\theta}$.

21. A tiene $+1$ o -1 en cada elemento en la diagonal; ocho posibilidades.

23. Las columnas de la matriz U de Fourier son vectores característicos de P porque $PU = \text{diag}(1, w, w^2, w^3)U$ (y $w = i$).

25. n^2 pasos para C directa multiplicada por x; sólo $n \log n$ pasos para F y F^{-1} por la TFR (y n para Λ).

27. $A^H A = \begin{bmatrix} 2 & 0 & 1+i \\ 0 & 2 & 1+i \\ 1-i & 1-i & 2 \end{bmatrix}$ y $AA^H = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ son matrices *hermitianas*. $(A^H A)^H = A^H A^{HH} = A^H A$ de nuevo.

29. cA sigue siendo hermitiana para c real; $(iA)^H = -iA^H = -iA$ es hermitiana sesgada.

31. $P^2 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $P^3 = I$, $P^{100} = P^{99}P = P$; λ = raíces cúbicas de $1 = 1$, $e^{2\pi i/3}$, $e^{4\pi i/3}$.

33. $C = \begin{bmatrix} 2 & 5 & 4 \\ 4 & 2 & 5 \\ 5 & 4 & 2 \end{bmatrix} = 2 + 5P + 4P^2$ tiene $\lambda(C) = \left\{ \begin{matrix} 2 + 5 + 4 \\ 2 + 5e^{2\pi i/3} + 4e^{4\pi i/3} \\ 2 + 5e^{4\pi i/3} + 4e^{8\pi i/3} \end{matrix} \right\}$.

35. $A = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & -1+i \\ 1+i & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1-i \\ -1-i & 1 \end{bmatrix}$.

$K = (iA^T) = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & -1-i \\ 1-i & 1 \end{bmatrix} \begin{bmatrix} 2i & 0 \\ 0 & -i \end{bmatrix} \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1+i \\ -1+i & 1 \end{bmatrix}$.

37. $V = \frac{1}{L}\begin{bmatrix}1+\sqrt{3} & -1+i\\ 1+i & 1+\sqrt{3}\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\frac{1}{L}\begin{bmatrix}1+\sqrt{3} & 1-i\\ -1-i & 1+\sqrt{3}\end{bmatrix}$ con $L^2 = 6+2\sqrt{3}$. $V = V^{\mathrm{H}}$ proporciona λ real, con una unitaria se obtiene $|\lambda| = 1$, de modo que con una traza cero se obtiene $\lambda = 1, -1$.

39. No multiplique e^{-ix} por e^{ix}; primero conjúguelos, luego $\int_0^{2\pi} e^{2ix}\,dx = [e^{2ix}/2i]_0^{2\pi} = 0$.

41. $R + iS = (R + iS)^{\mathrm{H}} = R^{\mathrm{T}} - iS^{\mathrm{T}}$; R es simétrica pero S es simétrica sesgada.

43. $[1]$ y $[-1]$; $\begin{bmatrix}a & b+ic\\ b-ic & -a\end{bmatrix}$ con $a^2+b^2+c^2 = 1$.

45. $(I - 2uu^{\mathrm{H}})^{\mathrm{H}} = I - 2uu^{\mathrm{H}}$; $(I - 2uu^{\mathrm{H}})^2 = I - 4uu^{\mathrm{H}} + 4u(u^{\mathrm{H}}u)u^{\mathrm{H}} = I$; la matriz uu^{H} proyecta sobre la recta que pasa por u.

47. Se tiene $A + iB = (A + iB)^{\mathrm{H}} = A^{\mathrm{T}} - iB^{\mathrm{T}}$. Así, $A = A^{\mathrm{T}}$ y $B = -B^{\mathrm{T}}$.

49. $A = \begin{bmatrix}1-i & 1-i\\ -1 & 2\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 4\end{bmatrix}\frac{1}{6}\begin{bmatrix}2+2i & -2\\ 1+i & 2\end{bmatrix} = S\Lambda S^{-1}$. valores característicos reales 1 y 4.

Conjunto de problemas 5.6, página 302

1. $C = N^{-1}BN = N^{-1}M^{-1}AMN = (MN)^{-1}A(MN)$; sólo $M^{-1}IM = I$ es semejante a I.

3. Si $\lambda_1, \ldots, \lambda_n$ son valores característicos de A, entonces $\lambda_1 + 1, \ldots, \lambda_n + 1$ son valores característicos de $A + I$. Por tanto, A y $A + I$ nunca tienen los mismos valores característicos, por lo que no pueden ser semejantes.

5. Si B es invertible, entonces $BA = B(AB)B^{-1}$ es semejante a AB.

7. El elemento (3, 1) de $M^{-1}AM$ es $g\cos\theta + h\operatorname{sen}\theta$, que es cero si $\tan\theta = -g/h$.

9. Los coeficientes son $c_1 = 1, c_2 = 2, d_1 = 1, d_2 = 1$; compruebe $Mc = d$.

11. La matriz reflexión con base v_1 y v_2 es $A = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$. La base V_1 y V_2 (¡misma reflexión!) proporciona $B = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$. Si $M = \begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$, entonces $A = MBM^{-1}$.

13. a) $D = \begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 2\\ 0 & 0 & 0\end{bmatrix}$. b) $D^3 = \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$ = matriz de terceras derivadas. Las terceras derivadas de 1, x y x^2 son cero, de modo que $D^3 = 0$. c) $\lambda = 0$ (triple); sólo un vector característico independiente $(1, 0, 0)$.

15. Estos valores característicos son $1, 1, 1, -1$. Matrices características $\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$, $\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$, $\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}$, $\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}$.

17. a) $TT^{\mathrm{H}} = U^{-1}AUU^{\mathrm{H}}A^{\mathrm{H}}(U^{-1})^{\mathrm{H}} = I$. b) Si T es triangular y unitaria, entonces sus elementos diagonales son cero porque las columnas deben ser vectores unitarios.

19. Los elementos 1, 1 de $T^{\mathrm{H}}T = TT^{\mathrm{H}}$ proporcionan $|t_{11}|^2 = |t_{11}|^2 + |t_{12}|^2 + |t_{13}|^2$, de modo que $t_{12} = t_{13} = 0$. Al comparar los elementos 2, 2 de $T^{\mathrm{H}}T = TT^{\mathrm{H}}$ se obtiene $t_{23} = 0$. Entonces, T debe ser diagonal.

21. Si $N = U\Lambda U^{-1}$, entonces $NN^{\mathrm{H}} = U\Lambda U^{-1}(U^{-1})^{\mathrm{H}}\Lambda^{\mathrm{H}}U^{\mathrm{H}}$ es igual a $U\Lambda\Lambda^{\mathrm{H}}U^{\mathrm{H}}$. Esto es lo mismo que $U\Lambda^{\mathrm{H}}\Lambda U^{\mathrm{H}} = (U\Lambda U^{-1})^{\mathrm{H}}(U\Lambda U^{-1}) = N^{\mathrm{H}}N$. Por tanto, N es normal.

23. Los valores característicos de $A(A - I)(A - 2I)$ son 0, 0, 0.

25. Siempre ¡$\begin{bmatrix}a^2+bc & ab+bd\\ ac+cd & bc+d^2\end{bmatrix} - (a+d)\begin{bmatrix}a & b\\ c & d\end{bmatrix} + (ad-bc)\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} = \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$!

27. $M^{-1}J_3M = 0$, por lo que las dos últimas desigualdades son fáciles. Al intentar para $MJ_1 = J_2M$ obliga a que la primera columna de M sea cero, por lo cual M no puede ser invertible. No es posible tener $J_1 = M^{-1}J_2M$.

29. $A^{10} = 2^{10}\begin{bmatrix}61 & 45\\ -80 & -59\end{bmatrix}$; $e^A = e^2\begin{bmatrix}13 & 9\\ -16 & -11\end{bmatrix}$.

31. $\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}$, $\begin{bmatrix}0 & 0\\ 1 & 1\end{bmatrix}$, $\begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}$, $\begin{bmatrix}0 & 1\\ 0 & 1\end{bmatrix}$ son semejantes; $\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$ de suyo y $\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$ de suyo.

33. a) $(M^{-1}AM)(M^{-1}x) = M^{-1}(Ax) = M^{-1}0 = 0$. b) Los espacios nulos de A y de $M^{-1}AM$ tienen la misma *dimensión*. Vectores y bases distintas.

35. $J^2 = \begin{bmatrix}c^2 & 2c\\ 0 & c^2\end{bmatrix}$, $J^3 = \begin{bmatrix}c^3 & 3c^2\\ 0 & c^3\end{bmatrix}$, $J^k = \begin{bmatrix}c^k & kc^{k-1}\\ 0 & c^k\end{bmatrix}$; $J^0 = I$, $J^{-1} = \begin{bmatrix}c^{-1} & -c^{-2}\\ 0 & c^{-1}\end{bmatrix}$.

37. $w(t) = \left(w(0) + tx(0) + \frac{1}{2}t^2y(0) + \frac{1}{6}t^3z(0)\right)e^{5t}$.

39. a) Se escoge M_i = matriz diagonal invertida para obtener $M_i^{-1}J_iM_i = M_i^{\mathrm{T}}$ en cada bloque. b) M_0 tiene estos bloques M_i en su diagonal para obtener $M_0^{-1}JM_0 = J^{\mathrm{T}}$. c) $A^{\mathrm{T}} = (M^{-1})^{\mathrm{T}}J^{\mathrm{T}}M^{\mathrm{T}}$ es $(M^{-1})^{\mathrm{T}}M_0^{-1}JM_0M^{\mathrm{T}} = (MM_0M^{\mathrm{T}})^{-1}A(MM_0M^{\mathrm{T}})$, y A^{T} es semejante a A.

41. a) Verdadero: Una tiene $\lambda = 0$, la otra no. b) Falso. Diagonalice una matriz no simétrica y Λ es simétrica. c) Falso $\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}$ y $\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$ son semejantes. d) Verdadero: todos los valores característicos de $A + I$ se incrementan por 1, con lo cual son distintos de los valores característicos de A.

43. Diagonales de 6 por 6 y de 4 por 4; AB tiene todos los mismos valores característicos que BA más $6 - 4$ ceros.

Conjunto de problemas 6.1, página 316

1. $ac - b^2 = 2 - 4 = -2 < 0$; $x^2 + 4xy + 2y^2 = (x + 2y)^2 - 2y^2$ (diferencia de cuadrados).

3. $\det(A - \lambda I) = \lambda^2 - (a + c)\lambda + ac - b^2 = 0$ proporciona $\lambda_1 = ((a + c) + \sqrt{(a - c)^2 + b^2})/2$ y $\lambda_2 = ((a + c) - \sqrt{(a - c)^2 + 4b^2})/2)$; $\lambda_1 > 0$ es una suma de números positivos; $\lambda_2 > 0$ porque $(a + c)^2 > (a - c)^2 + 4b^2$ se reduce a $ac > b^2$. Mejor método: el producto $\lambda_1\lambda_2 = ac - b^2$.

5. a) Positiva definida cuando $-3 < b < 3$.
b) $\begin{bmatrix}1 & b\\ b & 9\end{bmatrix} = \begin{bmatrix}1 & 0\\ b & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 9-b^2\end{bmatrix}\begin{bmatrix}1 & b\\ 0 & 1\end{bmatrix}$. c) El mínimo es $-\frac{1}{2(9-b^2)}$ cuando $\begin{bmatrix}1 & b\\ b & 9\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} = \begin{bmatrix}0\\ 1\end{bmatrix}$, que es $\begin{bmatrix}x\\ y\end{bmatrix} = \frac{1}{9-b^2}\begin{bmatrix}-b\\ 1\end{bmatrix}$. d) No hay mínimo; sean $y \to \infty$, $x = -3y$, entonces $x - y$ tiende a $-\infty$.

7. a) $A_1 = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}$ y $A_2 = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 2 & -2 \\ -1 & -2 & 11 \end{bmatrix}$.

b) $f_1 = (x_1 - x_2 - x_3)^2 = 0$ cuando $x_1 - x_2 - x_3 = 0$.

c) $f_2 = (x_1 - x_2 - x_3)^2 + (x_2 - 3x_3)^2 + x_3^2$; $L = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & -3 & 1 \end{bmatrix}$.

9. $A = \begin{bmatrix} 3 & 6 \\ 6 & 16 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$; los coeficientes de los cuadrados son los pivotes en D, mientras los coeficientes dentro de los cuadrados son columnas de L.

11. a) Los pivotes son a y $c - |b|^2/a$ y $\det A = ac - |b|^2$. b) Multiplique $|x_2|^2$ por $(c - |b|^2/a)$. c) Ahora, $x^H Ax$ es una suma de cuadrados. d) $\det = -1$ (indefinida) y $\det = +1$ (positiva definida).

13. $a > 1$ y $(a-1)(c-1) > b^2$. Esto significa que $A - I$ es positiva definida.

15. $f(x, y) = x^2 + 4xy + 9y^2 = (x + 2y)^2 + 5y^2$; $f(x, y) = x^2 + 6xy + 9y^2 = (x + 3y)^2$.

17. $x^T A^T Ax = (Ax)^T(Ax) =$ longitud al cuadrado $= 0$ sólo si $Ax = 0$. Debido a que A tiene columnas independientes, esto sólo ocurre cuando $x = 0$.

19. $A = \begin{bmatrix} 4 & -4 & 8 \\ -4 & 4 & -8 \\ 8 & -8 & 16 \end{bmatrix}$ sólo tiene un pivote $= 4$, rango $= 1$, valores característicos 24, 0, 0, $\det = A$.

21. $ax^2 + 2bxy + cy^2$ tiene un punto silla en $(0, 0)$ si $ac < b^2$. La matriz es *indefinida* ($\lambda < 0$ y $\lambda > 0$).

Conjunto de problemas 6.2, página 326

1. A es positiva definida para $a > 2$. B nunca es positiva definida: observe $\begin{bmatrix} 1 & 4 \\ 4 & 7 \end{bmatrix}$.

3. $\det A = -2b^3 - 3b^2 + 1$ es negativo en (y cerca de) $b = \frac{2}{3}$.

5. Si $x^T Ax > 0$ y $x^T Bx > 0$ para cualquier $x \neq 0$, entonces $x^T(A + B)x > 0$; condición (I).

7. λs positivos porque R es simétrica y $\sqrt{\Lambda} > 0$. $R = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$; $R = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$.

9. $|x^T Ay|^2 = |x^T R^T Ry|^2 = |(Rx)^T Ry|^2 \leq$ (por la desigualdad normal de Schwarz) $\|Rx\|^2 \|Ry\|^2 = (x^T R^T Rx)(y^T R^T Ry) = (x^T Ax)(y^T Ay)$.

11. $A = \begin{bmatrix} 3 & -\sqrt{2} \\ -\sqrt{2} & 2 \end{bmatrix}$ tiene $\lambda = 1$ y 4, ejes $1 \begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix}$ y $\frac{1}{2} \begin{bmatrix} \sqrt{2} \\ -1 \end{bmatrix}$ a lo largo de los vectores característicos.

13. *Matrices negativas definidas*: (I) $x^T Ax < 0$ para todos los vectores x diferentes de cero. (II) Todos los valores característicos de A que satisfacen $\lambda_i < 0$. (III) $\det A_1 < 0$, $\det A_2 > 0$, $\det A_3 < 0$. (IV) Todos los pivotes (sin intercambios de renglones) satisfacen $d_i < 0$. (V) Existe una matriz R con columnas independientes tal que $A = -R^T R$.

15. Falso (Q debe contener vectores característicos de A); verdadero (mismos valores característicos que A); verdadero ($Q^T AQ = Q^{-1}AQ$ es semejante a A); verdadero (los valores característicos de e^{-A} son $e^{-\lambda} > 0$).

17. Empiece desde $a_{jj} =$ (renglón j de R^T)(columna j de R) = longitud al cuadrado de la columna j de R. Así, $\det A = (\det R)^2 =$ (volumen del paralelepípedo R)$^2 \leq$ producto de las longitudes al cuadrado de todas las columnas de R. Este producto es $a_{11}a_{22} \cdots a_{nn}$.

19. $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ tiene pivotes $2, \frac{3}{2}, \frac{4}{3}$; $A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$ es singular; $A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

21. $x^T Ax$ no es positiva cuando $(x_1, x_2, x_3) = (0, 1, 0)$ debido al cero en la diagonal.

23. a) La condición de positiva definida requiere un determinante positivo (también que todas las $\lambda > 0$). b) todas las matrices proyección, excepto I, son singulares. c) Los elementos diagonales de D son sus valores característicos. d) La matriz negativa definida $-I$ tiene $\det = +1$ cuando n es par.

25. $\lambda_1 = 1/a^2$ y $\lambda_2 = 1/b^2$, de modo que $a = 1/\sqrt{\lambda_1}$ y $b = 1/\sqrt{\lambda_2}$. La elipse $9x^2 + 16y^2 = 1$ tiene ejes con semilongitudes $a = \frac{1}{3}$ y $b = \frac{1}{4}$.

27. $A = \begin{bmatrix} 9 & 3 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$; $C = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}$ tiene $CC^T = \begin{bmatrix} 4 & 8 \\ 8 & 25 \end{bmatrix}$.

29. $ax^2 + 2bxy + cy^2 = a\left(x + \frac{b}{a}y\right)^2 + \frac{ac-b^2}{a}y^2$; $2x^2 + 8xy + 10y^2 = 2(x + 2y)^2 + 2y^2$.

31. $x^T Ax = 2\left(x_1 - \frac{1}{2}x_2 - \frac{1}{2}x_3\right)^2 + \frac{3}{2}(x_2 - x_3)^2$; $x^T Bx = (x_1 + x_2 + x_3)^2$. B tiene un pivote.

33. A y $C^T AC$ tienen $\lambda_1 > 0$, $\lambda_2 = 0$. $C(t) = tQ + (1 - t)QR$, $Q = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $R = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$; C tiene un valor característico positivo y uno negativo, pero I tiene dos valores característicos positivos.

35. Los pivotes de $A - \frac{1}{2}I$ son 2.5, 5.9, −0.81, de modo que un valor característico de $A - \frac{1}{2}I$ es negativo. Por tanto, A tiene un valor característico menor que $\frac{1}{2}$.

37. rango$(C^T AC) \leq$ rango A, aunque también rango $(C^T AC) \geq$ rango $((C^T)^{-1}C^T ACC^{-1})$ = rango A.

39. No. Si C no es cuadrada, $C^T AC$ no es del mismo tamaño que la matriz A.

41. $\det \begin{bmatrix} 6 - 4\lambda/18 & -3 - \lambda/18 \\ -3 - \lambda/18 & 6 - 4\lambda/18 \end{bmatrix} = 0$ proporciona $\lambda_1 = 54$, $\lambda_2 = \dfrac{54}{5}$.

Vectores característicos $\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

43. *Grupos*: matrices ortogonales; e^{tA} para toda t; todas las matrices con $\det = 1$. Si A es positiva definida, el grupo de todas las potencias A^k contiene sólo matrices positivas definidas.

Conjunto de problemas 6.3, página 337

1. $A^TA = \begin{bmatrix} 5 & 20 \\ 20 & 80 \end{bmatrix}$ sólo tiene $\sigma_1^2 = 85$ con $v_1 = \begin{bmatrix} 1/\sqrt{17} \\ 4/\sqrt{17} \end{bmatrix}$, de modo que $v_2 = \begin{bmatrix} 4/\sqrt{17} \\ -1/\sqrt{17} \end{bmatrix}$.

3. $A^TA = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ tiene valores característicos $\sigma_1^2 = \dfrac{3+\sqrt{5}}{2}$ y $\sigma_2^2 = \dfrac{3-\sqrt{5}}{2}$. Como $A = A^T$, los vectores característicos de A^TA son los mismos que para A. Debido a que $\lambda_2 = \frac{1}{2}(1-\sqrt{5})$ es *negativa*, $\sigma_1 = \lambda_1$, pero $\sigma_2 = -\lambda_2$. Los vectores característicos unitarios son los mismos que en la sección 6.2 para A, excepto por el efecto de este signo menos (ya que se requiere $Av_2 = \sigma_2 u_2$):

$$u_1 = v_1 = \begin{bmatrix} \lambda_1/\sqrt{1+\lambda_1^2} \\ 1/\sqrt{1+\lambda_1^2} \end{bmatrix} \quad \text{y} \quad u_2 = -v_2 = \begin{bmatrix} \lambda_2/\sqrt{1+\lambda_2^2} \\ 1/\sqrt{1+\lambda_2^2} \end{bmatrix}.$$

5. $AA^T = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ tiene $\sigma_1^2 = 3$ con $u_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ y $\sigma_2^2 = 1$ con $u_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$.

$A^TA = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ tiene $\sigma_1^2 = 3$ con $v_1 = \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$, $\sigma_2^2 = 1$ con $v_2 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$,

y vector nulo, $v_3 = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$.

Así, $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = [u_1 \quad u_2] \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} [v_1 \quad v_2 \quad v_3]^T$.

7. $A = 12uv^T$ tiene un valor singular $\sigma_1 = 12$.

9. Multiplique $U\Sigma V^T$ usando columnas (de U) multiplicadas por renglones (de ΣV^T).

11. Para hacer singular a A, el cambio más pequeño iguala a cero su valor singular más pequeño, σ_2.

13. Los valores singulares de $A + I$ no son $\sigma_j + 1$. Provienen de valores característicos de $(A+I)^T(A+I)$.

15. $A^+ = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $B^+ = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$, $C^+ = \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \end{bmatrix}$.

A^+ es la inversa derecha de A; B^+ es la inversa izquierda de B.

17. $A^TA = \begin{bmatrix} 10 & 6 \\ 6 & 10 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 16 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$, tome las raíces cuadradas de 4 y 16 para obtener $S = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ y

$Q = AS^{-1} = \dfrac{1}{\sqrt{10}} \begin{bmatrix} 3 & 1 \\ -1 & 3 \end{bmatrix}$.

19. a) Con columnas independientes, el espacio renglón es todo $\mathbf{R}^n$; compruebe $(A^TA)A^+b = A^Tb$. b) $A^T(AA^T)^{-1}b$ está en el espacio renglón porque en este espacio se multiplica A^T por cualquier vector; así, $(A^TA)\,A^+b = A^TAA^T\,(AA^T)^{-1}b = A^Tb$. Con ambos casos se obtiene $A^TAx^+ = A^Tb$.

21. Tome $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$. Así, $AB = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$. A partir de C^+ en el problema 15 se tiene $A^+ = \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \end{bmatrix}$, $B^+ = \begin{bmatrix} 0 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix} = (AB)^+$, y $(AB)^+ \neq B^+A^+$.

23. $A = Q_1\Sigma Q_2^T \Rightarrow A^+ = Q_2\Sigma^+Q_1^T \Rightarrow AA^+ = Q_1\Sigma\Sigma^+Q_1^T$. Al elevar al cuadrado se obtiene $(AA^+)^2 = Q_1\Sigma\Sigma^+\Sigma\Sigma^+Q_1^T = Q_1\Sigma\Sigma^+Q_1^T$ y de manera semejante para $(AA^+)^2 = AA^+ = (AA^+)^T$ y A^+A proyectan sobre el espacio renglón y el espacio columna de A.

Conjunto de problemas 6.4, página 344

1. $P(x) = x_1^2 - x_1x_2 + x_2^2 - x_2x_3 + x_3^2 - 4x_1 - 4x_3$ tiene $\partial P/\partial x_1 = 2x_1 - x_2 - 4$, $\partial P/\partial x_2 = -x_1 + 2x_2 - x_3$, y $\partial P/\partial x_3 = -x_2 + 2x_3 - 4$.

3. Con $\partial P_1/\partial x = x + y = 0$ y $\partial P_1/\partial y = x + 2y - 3 = 0$ se obtiene $x = -3$ y $y = 3$. P_2 no tiene mínimo (sea $y \to \infty$). Está asociado con la matriz semidefinida $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

5. Escriba $x = (1, \ldots, 1)$ en el cociente de Rayleigh (el denominador se convierte en n): Debido a que $R(x)$ siempre está entre λ_1 y λ_n, se obtiene $n\lambda_1 \leq x^TAx =$ *suma de todas las* $a_{ij} \leq n\lambda_n$.

7. Debido a que $x^TBx > 0$ para todos los vectores x diferentes de cero, $x^T(A+B)x$ es mayor que x^TAx. Así, el cociente de Rayleigh es mayor para $A + B$ (de hecho, todos los n valores característicos se incrementan).

9. Debido a que $x^TBx > 0$, el cociente de Rayleigh es mayor para $A + B$ que el cociente para A.

11. Los valores característicos más pequeños en $Ax = \lambda x$ y $Ax = \lambda Mx$ son $\frac{1}{2}$ y $(3 - \sqrt{3})/4$.

13. a) $\lambda_j = \min_{S_j}[\max_{x \text{ en } S_j} R(x)] > 0$ significa que toda S_j contiene un vector x con $R(x) > 0$. b) $y = C^{-1}x$ proporciona el cociente $\overline{R}(y) = \dfrac{y^TC^TACy}{y^Ty} = \dfrac{x^TAx}{x^Tx} = R(x) > 0$.

15. El subespacio extremo S_2 es generado por los vectores característicos x_1 y x_2.

17. Si $Cx = C(A^{-1}b)$ es igual a d, entonces $CA^{-1}b - d$ es cero en el término de corrección en la ecuación (5).

Conjunto de problemas 6.5, página 350

1. $Ay = b$ es $4 \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 3/16 \\ 4/16 \\ 3/16 \end{bmatrix} = b = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$. El elemento lineal finito $U = \frac{3}{16}V_1 + \frac{4}{16}V_2 + \frac{3}{16}V_3$ es igual a la $u = \frac{3}{16}, \frac{4}{16}, \frac{3}{16}$ exacta en los nodos $x = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$.

3. $A_{33} = 3$, $b_3 = \frac{1}{3}$. Así, $A = 3 \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$, $b = \dfrac{1}{3} \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$,

$Ay = b$ proporciona $y = \dfrac{1}{9} \begin{bmatrix} 5 \\ 8 \\ 9 \end{bmatrix}$.

5. Integre por partes: $\int_0^1 -V_i''V_j\,dx = \int_0^1 V_i'V_j'\,dx - \left[V_i'V_j\right]_{x=0}^{x=1} = \int_0^1 V_i'V_j'\,dx =$ misma A_{ij}.

7. $A = 4$, $M = \frac{1}{3}$. La razón 12 (el cociente de Rayleigh sobre el subespacio de múltiplos de $V(x)$) es *mayor* que el valor característico verdadero $\lambda = \pi^2$.

9. La matriz masa M es $h/6$ veces la matriz tridiagonal 1, 4, 1.

Conjunto de problemas 7.2, página 357

1. Si Q es ortogonal, su norma es $\|Q\| = \text{máx}\,\|Qx\|/\|x\| = 1$ porque Q preserva la longitud: $\|Qx\| = \|x\|$ para toda x. Q^{-1} también es ortogonal y su norma es 1, de modo que $c(Q) = 1$.

3. $\|ABx\| \le \|A\|\|Bx\|$, por definición de la norma de A, y así $\|Bx\| \le \|B\|\|x\|$. Al dividir entre $\|x\|$ y maximizar, $\|AB\| \le \|A\|\|B\|$. Lo mismo se cumple para la inversa, $\|B^{-1}A^{-1}\| \le \|B^{-1}\|\|A^{-1}\|$; $c(AB) \le c(A)\,c(B)$ al multiplicar estas desigualdades.

5. En la definición $\|A\| = \text{máx}\,\|Ax\|/\|x\|$, x se escoge como el vector característico particular en cuestión: $\|Ax\| = |\lambda|\|x\|$, de modo que la razón es $|\lambda|$ y la razón *máxima* es *por lo menos* $|\lambda|$.

7. $A^{\mathrm{T}}A$ y AA^{T} tienen los mismos valores característicos, ya que con $A^{\mathrm{T}}Ax = \lambda x$ se obtiene $AA^{\mathrm{T}}(Ax) = A(A^{\mathrm{T}}Ax) = \lambda(Ax)$. La igualdad de los valores característicos más grandes significa $\|A\| = \|A^{\mathrm{T}}\|$.

9. $A = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$, $B = \begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}$, $\lambda_{\text{máx}}(A+B) > \lambda_{\text{máx}}(A) + \lambda_{\text{máx}}(B)$ (ya que $1 > 0 + 0$), y $\lambda_{\text{máx}}(AB) > \lambda_{\text{máx}}(A)\lambda_{\text{máx}}(B)$. Por tanto, $\lambda_{\text{máx}}(A)$ no es una norma.

11. a) Sí, $c(A) = \|A\|\|A^{-1}\| = c(A^{-1})$, ya que $(A^{-1})^{-1}$ es A de nuevo. b) $A^{-1}b = x$ conduce a $\dfrac{\|\delta b\|}{\|b\|} \le \|A\|\,\|A^{-1}\|\dfrac{\|\delta x\|}{\|x\|}$. Esto es $\dfrac{\|\delta x\|}{\|x\|} \ge \dfrac{1}{c}\dfrac{\|\delta b\|}{\|b\|}$.

13. $\|A\| = 2$ y $c = 1$; $\|A\| = \sqrt{2}$ y c es infinita (¡singular!); $\|A\| = \sqrt{2}$ y $c = 1$.

15. Si $\lambda_{\text{máx}} = \lambda_{\text{mín}} = 1$, entonces todos los $\lambda_i = 1$ y $A = SIS^{-1} = I$. Las únicas matrices con $\|A\| = \|A^{-1}\| = 1$ son *matrices ortogonales*, porque $A^{\mathrm{T}}A$ tiene que ser I.

17. El residuo $b - Ay = (10^{-7}, 0)$ es mucho menor que $b - Az = (0.0013, 0.0016)$. Sin embargo, z está mucho más cerca de la solución que y.

19. $x_1^2 + \cdots + x_n^2$ no es más pequeño que $\text{máx}(x_i^2) = (\|x\|_\infty)^2$ y no es más grande que $(|x_1| + \cdots + |x_n|)^2$, que es $(\|x\|_1)^2$. Ciertamente, $x_1^2 + \cdots + x_n^2 \le n\,\text{máx}(x_i^2)$, de modo que $\|x\| \le \sqrt{n}\|x\|_\infty$. Se escoge $y = (\operatorname{sign} x_1, \operatorname{sign} x_2, \ldots, \operatorname{sign} x_n)$ para obtener $x \cdot y = \|x\|_1$. Por la desigualdad de Schwarz, esto es cuando mucho $\|x\|\|y\| = \sqrt{n}\|x\|$. Se escoge $x = (1, 1, \ldots, 1)$ para razones máximas $\sqrt{n}$.

21. La inversa exacta de la matriz de Hilbert de 3 por 3 es $A^{-1} = \begin{bmatrix}9 & -36 & 30\\ -36 & 192 & -180\\ 30 & -180 & 180\end{bmatrix}$.

23. El $\|x\| = \|A^{-1}b\|$ más grande es $1/\lambda_{\text{mín}}$; el error más grande es $10^{-16}/\lambda_{\text{mín}}$.

25. Intercambie $\begin{bmatrix}1 & 0\\ 2 & 2\end{bmatrix}$ a $\begin{bmatrix}2 & 2\\ 1 & 0\end{bmatrix} \to \begin{bmatrix}2 & 2\\ 0 & -1\end{bmatrix} = U$ con $P = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$ y

$L = \begin{bmatrix}1 & 0\\ 0.5 & 1\end{bmatrix}$. $A \to \begin{bmatrix}2 & 2 & 0\\ 1 & 0 & 1\\ 0 & 2 & 0\end{bmatrix} \to \begin{bmatrix}2 & 2 & 0\\ 0 & -1 & 1\\ 0 & 2 & 0\end{bmatrix} \to \begin{bmatrix}2 & 2 & 0\\ 0 & 2 & 0\\ 0 & -1 & 1\end{bmatrix} \to$ $\begin{bmatrix}2 & 2 & 0\\ 0 & 2 & 0\\ 0 & 0 & 1\end{bmatrix} = U$. Entonces $PA = LU$ con $P = \begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{bmatrix}$ y $L = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0.5 & -0.5 & 1\end{bmatrix}$.

Conjunto de problemas 7.3, página 365

1. $u_0 = \begin{bmatrix}1\\ 0\end{bmatrix}$, $u_1 = \begin{bmatrix}2\\ -1\end{bmatrix}$, $u_2 = \begin{bmatrix}5\\ -4\end{bmatrix}$, $u_3 = \begin{bmatrix}14\\ -13\end{bmatrix}$; $u_\infty = \dfrac{1}{\sqrt{2}}\begin{bmatrix}1\\ -1\end{bmatrix}$ normalizado a un vector unitario.

3. $u_k/\lambda_1^k = c_1x_1 + c_2x_2(\lambda_2/\lambda_1)^k + \cdots + c_nx_n(\lambda_n/\lambda_1)^k \to c_1x_1$ si todas las razones $|\lambda_i/\lambda_1| < 1$. La razón máxima controla, cuando k es grande. $A = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$ tiene $|\lambda_2| = |\lambda_1|$ y ninguna convergencia.

5. $Hx = x - (x-y)\dfrac{2(x-y)^{\mathrm{T}}x}{(x-y)^{\mathrm{T}}(x-y)} = x - (x-y) = y$. Entonces $H(Hx) = Hy$ es $x = Hy$.

7. $U = \begin{bmatrix}1 & 0\\ 0 & H\end{bmatrix} \begin{bmatrix}1 & 0 & 0\\ 0 & -\frac{3}{5} & -\frac{4}{5}\\ 0 & -\frac{4}{5} & \frac{3}{5}\end{bmatrix} = U^{-1}$ y entonces $U^{-1}AU = \begin{bmatrix}1 & -5 & 0\\ -5 & \frac{9}{25} & \frac{12}{25}\\ 0 & \frac{12}{25} & \frac{16}{25}\end{bmatrix}$.

9. $\begin{bmatrix}\cos\theta & \operatorname{sen}\theta\\ \operatorname{sen}\theta & 0\end{bmatrix} = QR = \begin{bmatrix}\cos\theta & -\operatorname{sen}\theta\\ \operatorname{sen}\theta & \cos\theta\end{bmatrix}\begin{bmatrix}1 & \cos\theta\operatorname{sen}\theta\\ 0 & -\operatorname{sen}^2\theta\end{bmatrix}$.

Entonces $RQ = \begin{bmatrix}c(1+s^2) & -s^3\\ -s^3 & -s^2c\end{bmatrix}$.

11. *Suponga* que $(Q_0 \cdots Q_{k-1})(R_{k-1} \cdots R_0)$ es la factorización QR de A^k (lo cual ciertamente es verdadero si $k = 1$). Por construcción, $A_{k+1} = R_kQ_k$ de modo que $R_k = A_{k+1}Q_k^{\mathrm{T}} = (Q_k^{\mathrm{T}} \cdots Q_0^{\mathrm{T}}AQ_0 \cdots Q_k)Q_k^{\mathrm{T}}$. Al multiplicar por la derecha por $(R_{k-1} \cdots R_0)$, la hipótesis proporciona $R_k \cdots R_0 = Q_k^{\mathrm{T}} \cdots Q_0^{\mathrm{T}}A^{k+1}$. Después de pasar las Qs al miembro izquierdo, este es el resultado requerido para A^{k+1}.

13. A tiene valores característicos 4 y 2. Escriba un vector unitario en el renglón 1 de P: es ya sea $\dfrac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}$ y $PAP^{-1} = \begin{bmatrix}2 & -4\\ 0 & 4\end{bmatrix}$ o $\dfrac{1}{\sqrt{10}}\begin{bmatrix}1 & -3\\ 3 & 1\end{bmatrix}$ y $PAP^{-1} = \begin{bmatrix}4 & -4\\ 0 & 2\end{bmatrix}$.

15. P_{ij} usa $4n$ multiplicaciones (2 por cada elemento en los renglones i y j). Al factorizar $\cos\theta$, los elementos 1 y $\pm\tan\theta$ sólo requieren $2n$ multiplicaciones, lo cual lleva a $\frac{2}{3}n^3$ para PR.

Conjunto de problemas 7.4, página 372

1. $D^{-1}(-L-U) = \begin{bmatrix}0 & \frac{1}{2} & 0\\ \frac{1}{2} & 0 & \frac{1}{2}\\ 0 & \frac{1}{2} & 0\end{bmatrix}$, valores característicos $\mu = 0, \pm 1/\sqrt{2}$; $(D+L)^{-1}$

$(-U) = \begin{bmatrix}1 & \frac{1}{2} & 0\\ 0 & \frac{1}{4} & \frac{1}{2}\\ 0 & \frac{1}{8} & \frac{1}{4}\end{bmatrix}$, valores característicos 0, 0, $1/2$; $\omega_{\text{ópt}} = 4 - 2\sqrt{2}$, reduciendo $\lambda_{\text{máx}}$ a $3 - 2\sqrt{2} \approx 0.2$.

3. $Ax_k = (2 - 2\cos k\pi h)x_k$; $Jx_k = \frac{1}{2}(\text{sen}\, 2k\pi h, \text{sen}\, 3k\pi h + \text{sen}\, k\pi h, \ldots) =$ $(\cos k\pi h)x_k$. Para $h = \dfrac{1}{4}$, A tiene valores característicos $2 - 2\cos\dfrac{\pi}{4} = 2 - \sqrt{2}$, $2 - \cos\dfrac{\pi}{2} = 2$, $2 - \cos\dfrac{3\pi}{4} = 2 + \sqrt{2}$.

5. $J = D^{-1}(L + U) = -\begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & \frac{1}{4} \\ \frac{2}{5} & \frac{2}{5} & 0 \end{bmatrix}$; los tres círculos tienen radios $r_1 = \dfrac{2}{3}$, $r_2 = \dfrac{1}{4}$, $r_3 = \dfrac{4}{5}$. Sus centros están en cero, de modo que todo $|\lambda_i| \leq 4/5 < 1$.

7. $-D^{-1}(L + U) = \begin{bmatrix} 0 & -b/a \\ -c/d & 0 \end{bmatrix}$ tiene $\mu = \pm\left(\dfrac{bc}{ad}\right)^{1/2}$; $-(D + L)^{-1}U =$ $\begin{bmatrix} 0 & -b/a \\ 0 & bc/ad \end{bmatrix}$, $\lambda = 0, bc/ad$; $\lambda_{\text{máx}}$ es igual a $\mu^2_{\text{máx}}$.

9. Si $Ax = \lambda x$, entonces $(I - A)x = (1 - \lambda)x$. Los valores característicos reales de $B = I - A$ tienen $|1 - \lambda| < 1$, en el supuesto de que λ esté entre 0 y 2.

11. Siempre $\|AB\| \leq \|A\|\|B\|$. Se escoge $A = B$ para encontrar $\|B^2\| \leq \|B\|^2$. Luego se es- coge $A = B^2$ para encontrar $\|B^3\| \leq \|B^2\|\|B\| \leq \|B^3\|$. Continúe (o aplique inducción). Debido a que $\|B\| \geq$ máx $|\lambda(B)|$, no es sorprendente que con $\|B\| < 1$ se obtenga convergencia.

13. Jacobi tiene $S^{-1}T = \frac{1}{3}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ con $|\lambda|_{\text{máx}} = \frac{1}{3}$. Gauss-Seidel tiene $S^{-1}T = \begin{bmatrix} 0 & \frac{1}{3} \\ 0 & \frac{1}{9} \end{bmatrix}$ con $|\lambda|_{\text{máx}} = \frac{1}{9} = (|\lambda|_{\text{máx}}$ para Jacobi$)^2$.

15. Sobrerrelajamiento (SRS) sucesivo en MATLAB.

17. Todas las sumas máximas de los renglones son $|\lambda| \leq 0.9$ y $|\lambda| \leq 4$. Los círculos alre- dedor de los elementos de la diagonal proporcionan cotas más estrechas. Primera A: el círculo $|\lambda - 0.2| \leq 0.7$ contiene a los demás círculos $|\lambda - 0.3| \leq 0.5$ y $|\lambda - 0.1| \leq$ 0.6 y a todos los tres valores característicos. Segunda A: el círculo $|\lambda - 2| \leq 2$ contie- ne al círculo $|\lambda - 2| \leq 1$, a todos los tres valores característicos $2 + \sqrt{2}$, 2, y $2 - \sqrt{2}$.

19. $r_1 = b - \alpha_1 Ab = b - (b^{\mathrm{T}}b/b^{\mathrm{T}}Ab)Ab$ es ortogonal a $r_0 = b$: *los residuos* $r = b$ *$- Ax$ son ortogonales en cada paso*. Para demostrar que p_1 es ortogonal a $Ap_0 = Ab$, p_1 se simplifica a cP_1: $P_1 = \|Ab\|^2 b - (b^{\mathrm{T}}Ab)Ab$ y $c = b^{\mathrm{T}}b/(b^{\mathrm{T}}Ab)^2$. Ciertamen- te $(Ab)^{\mathrm{T}}P_1 = 0$, ya que $A^{\mathrm{T}} = A$. (Esta simplificación coloca a α_1 en $p_1 = b - \alpha_1 Ab +$ $(b^{\mathrm{T}}b - 2\alpha_1 b^{\mathrm{T}}Ab + \alpha_1^2\|Ab\|^2)b/b^{\mathrm{T}}b$. Para un buen análisis, consulte la obra *Numeri- cal Linear Algebra* de Trefethen y Bau).

Conjunto de problemas 8.1, página 381

1. Los vértices están en (0, 6), (2, 2), (6, 0); consulte la figura 8.3.

3. Las restricciones proporcionan $3(2x + 5y) + 2(-3x + 8y) \leq 9 - 10$, o bien, $31y \leq -1$. No es posible tener $y \geq 0$.

5. $x \geq 0$, $y \geq 0$, con restricciones adicionales de que $x + y \leq 0$ sólo admite el punto (0, 0).

7. x (bonos al 5%) $= z$ (bonos al 9%) $= 20\,000$ y y (bonos al 6%) $= 60\,000$.

9. El costo a minimizar es $1000x + 2000y + 3000z + 1500u + 3000v + 3700w$. Las cantidades x, y, z a Chicago y u, v, w a Nueva Inglaterra satisfacen $x + u = 1\,000\,000$; $y + v = 1\,000\,000$; $z + w = 1\,000\,000$; $x + y + z = 800\,000$; $u + v + w = 2\,200\,000$.

Conjunto de problemas 8.2, página 391

1. En el presente, $x_4 = 4$ y $x_5 = 2$ están en la base, y el costo es cero. La variable de entrada debe ser x_3 para reducir el costo. La variable de salida debe ser x_5, ya que 2/1 es menor que 4/1. Con x_3 y x_4 en la base, las restricciones proporcionan $x_3 = 2$, $x_4 = 2$, y ahora el costo es $x_1 + x_2 - x_3 = -2$.

3. Los "costos reducidos" son $r = [1 \quad 1]$, de modo que el cambio no es bueno y el vértice es óptimo.

5. En P, $r = [-5 \quad 3]$; así en Q, $r = [\frac{5}{3} \quad -\frac{1}{3}]$; R es óptimo porque $r \geq 0$.

7. Para un problema de maximización, la prueba de detención se convierte en $r \leq 0$. Si fracasa, y el i-ésimo componente es el más grande, entonces esa columna de N entra en la base; la regla **8C** para el vector que sale de la base es la misma.

9. $BE = B[\cdots v \cdots] = [\cdots u \cdots]$, ya que $Bv = u$. Por tanto, la matriz correcta es E.

11. Si $Ax = 0$, entonces $Px = x - A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}Ax = x$.

Conjunto de problemas 8.3 página 399

1. Maximizar $4y_1 + 11y_2$, con $y_1 \geq 0$, $y_2 \geq 0$, $2y_1 + y_2 \leq 1$, $3y_2 \leq 1$; el original tiene $x_1^* = 2$, $x_2^* = 3$, el dual tiene $y_1^* = \frac{1}{3}$, $y_2^* = \frac{1}{3}$, *costo* $= 5$.

3. El dual maximiza yb, con $y \geq c$. En consecuencia, $x = b$ y $y = c$ son factibles, y proporcionan el mismo valor cb para el costo en el original y en el dual; por **8F** deben ser óptimas. Si $b_1 < 0$, entonces la óptima x^* cambia a $(0, b_2, \ldots, b_n)$ y $y^* = (0, c_2, \ldots, c_n)$.

5. $b = [0 \quad 1]^{\mathrm{T}}$, y $c = [-1 \quad 0]$.

7. Debido a que $cx = 3 = yb$, x y y por **8F** son óptimas.

9. $x^* = [1 \quad 0]^{\mathrm{T}}$ y $y^* = [1 \quad 0]$, con $y^*b = 1 = cx^*$. Las segundas desigualdades tanto en $Ax^* \geq b$ como en $y^*A \leq c$ son estrictas, por lo que las segundas componentes de y^* y x^* son cero.

11. a) $x_1^* = 0$, $x_2^* = 1$, $x_3^* = 0$, $c^{\mathrm{T}}x = 3$. b) Es el primer cuadrante con el tetraedro en el vértice de corte. c) Maximizar y_1, sujeto a $y_1 \geq 0$, $y_1 \leq 5$, $y_1 \leq 3$, $y_1 \leq 4$; $y_1^* = 3$.

13. Aquí $c = [1 \quad 1 \quad 1]$ con $A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. No hay restricción $x \geq 0$, por lo que el dual tiene la igualdad $yA = c$ (o $A^{\mathrm{T}}y = c^{\mathrm{T}}$). Así se obtiene $2y_1 = 1$ y $y_1 = 1$ y $y_2 = 2$ y *ninguna solución factible*. Así, el máximo del original debe ser ∞: $x_1 = -N$ y $x_2 = 2N$ y $x_3 = 0$ proporcionan Costo $= x_1 + x_2 + x_3 = N$ (arbitrariamente grande).

15. Las columnas de $\begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix}$ o $\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$.

17. Tome $y = [1 \quad -1]$; entonces $yA \geq 0$, $yb < 0$.

Conjunto de problemas 8.4, página 406

1. El flujo maximal es 13, con el corte minimal que separa el nodo 6 de los otros nodos.
3. Al incrementar la capacidad de los tubos del nodo 4 al nodo 6 o del nodo 4 al nodo 5 se obtiene el incremento máximo en el flujo maximal. El flujo maximal crece de 8 a 9.
5. Asigne capacidades = 1 a todas las aristas. Entonces, el número máximo de rutas ajenas desde s hasta t es igual al flujo máximo. El número mínimo de aristas cuya eliminación desconecta s de t es el corte mínimo. Así, flujo máx = corte mín.
7. Los renglones 1, 4 y 5 violan la condición de Hall; la submatriz de 3 por 3 que proviene de los renglones 1, 4 y 5 y de las columnas 1, 2, 5 tiene $3 + 3 > 5$.
9. a) La matriz tiene $2n$ unos, que no es posible cubrir por menos de n rectas porque cada una cubre exactamente dos unos. Se requieren n rectas; debe haber un apareamiento completo.

   b) $\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$. Los unos pueden cubrirse con cuatro rectas; cinco matrimonios no es posible.
11. Si cada $m + 1$ se casa sólo con el hombre aceptable m, entonces no hay nadie que se case con el #1 (aun cuando todos sean aceptables para el #1).
13. Con el algoritmo 1 se obtiene 1-3, 3-2, 2-5, 2-4, 4-6, y con el algoritmo 2, 2-5, 4-6, 2-4, 3-2, 1-3. Estos son los árboles generadores más cortos de la misma longitud.
15. a) Los renglones 1, 3 y 5 sólo tienen unos en las columnas 2 y 4. b) Las columnas 1, 3 y 5 (en los renglones 2 y 4). c) La submatriz cero desde los renglones 1, 3 y 5 y las columnas 1, 3 y 5. d) Los renglones 2 y 4 y las columnas 2 y 4 cubren todos los unos.

Conjunto de problemas 8.5, página 413

1. $-10x_1 + 70(1 - x_1) = 10x_1 - 10(1 - x_1)$, o bien, $x_1 = \frac{4}{5}, x_2 = \frac{1}{5}$; $-10y_1 + 10(1 - y_1) = 70y_1 - 10(1 - y_1)$, o bien, $y_1 = \frac{1}{5}$, $y_2 = \frac{4}{5}$; pago medio $yAx = 6$.
3. Si X escoge la columna j, Y elegirá su elemento más pequeño a_{ij} (en el renglón i). X no se moverá, porque este es el mayor elemento en ese renglón. En el problema 2, $a_{12} = 2$ era un equilibrio de este tipo. Si se intercambian el 2 y el 4 debajo del a_{12}, ningún elemento tiene esta propiedad, por lo que se requieren estrategias mezcladas.
5. La mejor estrategia para X combina las dos rectas para obtener una recta horizontal, garantizando esta altura de $7/3$. La combinación es $\frac{2}{3}(3y + 2(1 - y)) + \frac{1}{3}(y + 3(1 - y)) = 7/3$, por lo que X escoge las columnas con frecuencias $\frac{2}{3}, 0, \frac{1}{3}$.
7. Para las columnas, se quiere $x_1 a + (1 - x_1)b = x_1 c + (1 - x_1)d = u$, de modo que $x_1(a - b - c + d) = d - b$. Para los renglones, $y_1 a + (1 - y_1)c = y_1 b + (1 - y_1)d = v$ intercambia b y c. Compare u con v:

   $u = x_1(a - b) + b = \dfrac{(a-b)(d-b)}{a-b-c+d} + b = \dfrac{ad - bc}{a-b-c+d}$ = lo mismo después de $b \leftrightarrow c = v$.

9. El máximo interior es el mayor de y_1 y y_2; x se concentra en ése. Sujeto a $y_1 + y_2 = 1$, el mínimo de la *y más grande* es $\frac{1}{2}$. Observe $A = I$.
11. $Ax^* = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}^T$ y $yAx^* = \frac{1}{2}y_1 + \frac{1}{2}y_2 = \frac{1}{2}$ para todas las estrategias de Y; $y^*A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -1 & -1 \end{bmatrix}$ y $y^*Ax = \frac{1}{2}x_1 + \frac{1}{2}x_2 - x_3 - x_4$, que no pueden exceder $\frac{1}{2}$; en medio está $y^*Ax^* = \frac{1}{2}$.
13. Valor 0 (juego justo). X escoge 2 o 3, y escoge impar o par: $x^* = y^* = \left(\frac{1}{2}, \frac{1}{2}\right)$.

Conjunto de problemas A, página 420

1. a) Mayor dim $(\mathbf{S} \cap \mathbf{T}) = 7$ cuando $\mathbf{S} \subset \mathbf{T}$. b) Menor dim $(\mathbf{S} \cap \mathbf{T}) = 2$. c) Menor dim $(\mathbf{S} + \mathbf{T}) = 8$ cuando $\mathbf{S} \subset \mathbf{T}$. d) Mayor dim $(\mathbf{S} + \mathbf{T}) = 13$ (todo $\mathbf{R}^{13}$).
3. $\mathbf{V} + \mathbf{W}$ y $\mathbf{V} \cap \mathbf{W}$ contienen $\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ 0 & a_{32} & a_{33} & a_{34} \\ 0 & 0 & a_{43} & a_{44} \end{bmatrix}$ y $\begin{bmatrix} a_{11} & a_{12} & 0 & 0 \\ 0 & a_{22} & a_{23} & 0 \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}$.

   $\dim(\mathbf{V} + \mathbf{W}) = 13$ y $\dim(\mathbf{V} \cap \mathbf{W}) = 7$; se suma para obtener $20 = \dim \mathbf{V} + \dim \mathbf{W}$.
5. Las rectas que pasan por (1, 1, 1) y (1, 1, 2) tienen $\mathbf{V} \cap \mathbf{W} = \{0\}$.
7. Una base para $\mathbf{V} + \mathbf{W}$ es v_1, v_2, w_1; dim $(\mathbf{V} \cap \mathbf{W}) = 1$ con base $(0, 1, -1, 0)$.
9. La intersección de los espacios columna es la recta que pasa por $y = (6, 3, 6)$:

   $y = \begin{bmatrix} 1 & 5 \\ 3 & 0 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ coincide con $[A \quad B]x = \begin{bmatrix} 1 & 5 & 3 & 0 \\ 3 & 0 & 0 & 1 \\ 2 & 4 & 0 & 2 \end{bmatrix} \begin{vmatrix} 1 \\ 1 \\ -2 \\ -3 \end{vmatrix} = 0.$

   La dimensión de los espacios columna es 2. La suma y la intersección de ellos son $3 + 1 = 2 + 2$.
11. $F_2 \otimes F_2 = \begin{bmatrix} F_2 & F_2 \\ F_2 & -F_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$.
13. $A_{3D} = (A_{1D} \otimes I \otimes I) + (I \otimes A_{1D} \otimes I) + (I \otimes I \otimes A_{1D})$.

Conjunto de problemas B, página 427

1. $J = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ (A es diagonalizable); $J = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (vectores característicos (1, 0, 0) y (2, −1, 0).
3. $e^{Bt} = \begin{bmatrix} 1 & t & 2t \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = I + Bt$ porque $B^2 = 0$. También $e^{Jt} = I + Jt$.
5. $J = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix}$ (valores característicos distintos); $J = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ (B tiene $\lambda = 0, 0$, pero rango 1).



Factorizaciones matriciales

1. $A = LU = \begin{pmatrix} L \text{ triangular inferior} \\ \text{unos en la diagonal} \end{pmatrix} \begin{pmatrix} U \text{ triangular superior} \\ \text{pivotes en la diagonal} \end{pmatrix}$

   Requerimientos: Ningún intercambio de renglones mientras la eliminación gaussiana reduce A a U.

2. $A = LDU = \begin{pmatrix} L \text{ triangular inferior} \\ \text{unos en la diagonal} \end{pmatrix} \begin{pmatrix} \text{matriz de pivotes} \\ D \text{ es diagonal} \end{pmatrix} \begin{pmatrix} U \text{ triangular superior} \\ \text{unos en la diagonal} \end{pmatrix}$

   Requerimientos: Ningún intercambio de renglones. Los pivotes de D se dividen para dejar unos en U. Si A es simétrica, entonces U es L^{T} y $A = LDL^{\mathrm{T}}$.

3. $PA = LU$ (matriz permutación para evitar ceros en las posiciones pivote).

   Requerimientos: A es invertible: Así, P, L, U son invertibles. P realiza el intercambio de renglones de antemano. Alternativa: $A = L_1P_1U_1$.

4. $EA = R$ (E invertible de m por m)(cualquier A) = rref(A).

   Requerimientos: ¡Ninguno! *La forma escalonada reducida* R tiene r renglones pivote y columnas pivote. El único elemento diferente de cero en una columna es el pivote unitario. Los $m - r$ últimos renglones de E son una base del espacio nulo izquierdo de A, y así las r primeras columnas de E^{-1} son una base para el espacio columna de A.

5. $A = CC^{\mathrm{T}}$ = (Matriz triangular inferior C)(la traspuesta es triangular superior).

   Requerimientos: A es simétrica y positiva definida (todos los n pivotes en D son positivos). La *factorización de Cholesky* tiene $C = L\sqrt{D}$.

6. $A = QR$ = (columnas ortonormales en Q)(triangular superior R).

   Requerimientos: A tiene columnas independientes. Éstas son *ortogonalizadas* en Q mediante el proceso de Gram-Schmidt. Si A es cuadrada, entonces $Q^{-1} = Q^{\mathrm{T}}$.

7. $A = S\Lambda S^{-1}$ = (vectores característicos en S)(valores característicos en Λ)(vectores característicos izquierdos en S^{-1}).

   Requerimientos: A debe tener n vectores característicos linealmente independientes.

8. $A = Q\Lambda Q^{\mathrm{T}}$ = (matriz ortogonal Q)(matriz Λ de valores característicos reales)(Q^{T} es Q^{-1}).

   Requerimientos: A es *simétrica*. Este es el teorema espectral.

9. $A = MJM^{-1}$ = (vectores característicos generalizados en M)(bloques de Jordan en J)(M^{-1}).

   Requerimientos: A es cualquier matriz cuadrada. La *forma de Jordan* J tiene un bloque para cada vector característico de A independiente. Cada bloque tiene un valor característico.

10. $A = U\Sigma V^{\mathrm{T}} = \begin{pmatrix} U \text{ ortogonal} \\ \text{es de } m \times m \end{pmatrix} \begin{pmatrix} \text{Matriz } \Sigma \text{ de } m \times n \\ \sigma_1, \ldots, \sigma_r \text{ en la diagonal} \end{pmatrix} \begin{pmatrix} V \text{ ortogonal} \\ \text{es de } n \times n \end{pmatrix}$

   Requerimientos: Ninguno. Esta *descomposición del valor singular* (DVS) tiene los vectores característicos de AA^{T} en U y los de $A^{\mathrm{T}}A$ en V; $\sigma_i = \sqrt{\lambda_i(A^{\mathrm{T}}A)} = \sqrt{\lambda_i(AA^{\mathrm{T}})}$.

11. $A^+ = V\Sigma^+U^{\mathrm{T}} = \begin{pmatrix} \text{ortogonal} \\ \text{de } n \text{ por } n \end{pmatrix} \begin{pmatrix} \text{diagonal } 1/\sigma_1, \ldots, 1/\sigma_r \\ \text{de } n \text{ por } m \end{pmatrix} \begin{pmatrix} \text{ortogonal} \\ \text{de } m \text{ por } m \end{pmatrix}$

    Requerimientos: Ninguno. La *seudoinversa* tiene A^+A = proyección sobre el espacio renglón de A y AA^+ = proyección sobre el espacio columna. La solución más corta por mínimos cuadrados de $Ax = b$ es $\hat{x} = A^+b$. Esto resuelve $A^{\mathrm{T}}A\hat{x} = A^{\mathrm{T}}b$.

12. $A = QH$ = (matriz ortogonal Q) matriz H positiva definida simétrica).

    Requerimientos: A es invertible. Esta *descomposición polar* tiene $H^2 = A^{\mathrm{T}}A$. El factor H es semidefinido si A es singular. La descomposición polar inversa $A = KQ$ tiene $K^2 = AA^{\mathrm{T}}$. Con base en la DVS, ambas tienen $Q = UV^{\mathrm{T}}$.

13. $A = U\Lambda U^{-1}$ = (U unitaria)(matriz valor característico Λ)($U^{-1} = U^{\mathrm{H}} = \overline{U}^{\mathrm{T}}$).

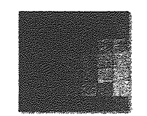
    Requerimientos: A *es normal*: $A^{\mathrm{H}}A = AA^{\mathrm{H}}$. Sus vectores característicos ortonormales (y tal vez complejos) son las columnas de U. Los λs son complejos a menos que $A = A^{\mathrm{H}}$.

14. $A = UTU^{-1}$ = (U unitaria)(T triangular con λs en la diagonal)($U^{-1} = U^{\mathrm{H}}$).

    Requerimientos: *Triangularización de Schur* de cualquier A cuadrada. Existe una matriz U con columnas ortonormales que hace triangular a $U^{-1}AU$.

15. $F_n = \begin{bmatrix} I & D \\ I & -D \end{bmatrix} \begin{bmatrix} F_{n/2} & \\ & F_{n/2} \end{bmatrix} \begin{bmatrix} \text{Permutación} \\ \text{par-impar} \end{bmatrix}$ = un paso de la **TFR**.

    Requerimientos: F_n = matriz de Fourier con elementos w^{jk} donde $w^n = 1$, $w = e^{2\pi i/n}$. Así, $F_n\overline{F}_n = nI$. D tiene $1, w, w^2, \ldots$ en su diagonal. Para $n = 2^\ell$, la *transformada de Fourier rápida* (TFR) tiene $\frac{1}{2}n\ell$ multiplicaciones de las ℓ etapas de las Ds.



Glosario: Un diccionario de álgebra lineal

Base de V Vectores independientes $v_1, \ldots, v_d$ cuya combinación lineal proporciona todo v en $\mathbf{V}$. ¡Un espacio vectorial tiene muchas bases!

Base normal de $\mathbf{R}^n$ Columnas de la matriz identidad de n por n (se escribe i, j, k, en $\mathbf{R}^3$).

Cociente de Rayleigh $q(x) = x^TAx/x^Tx$ Para cada $A = A^T$, $\lambda_{mín}\ \Sigma\ q(x)\ \Sigma\ \lambda_{máx}$. Estos extremos se alcanzan en los vectores característicos x para $\lambda_{mín}(A)$ y $\lambda_{máx}(A)$.

Cofactor C_{ij} Quitar el renglón i y la columna j; multiplicar el determinante por $(-1)^{i+j}$.

Columnas pivote de A Columnas que contienen pivotes después de la reducción de renglones; no son combinaciones de columnas anteriores. Las columnas pivote son una base del espacio columna.

Complemento de Schur $S = D - CA^{-1}B$ Aparece en la eliminación por bloques de $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$.

Compare $L_0 = 2$ con los números de Fibonacci.

Combinación lineal $cv + dw$ o Σc_jv_j Suma vectorial y multiplicación por un escalar.

Conjugado complejo $\bar{z} = a - ib$ para cualquier número complejo $z = a + ib$. Así, $z\bar{z} = |z|^2$.

Conjunto $v_1, \ldots, v_m$ generador de V Todo vector en V es una combinación de $v_1, \ldots, v_m$.

Cuatro subespacios fundamentales de A $C(A)$, $N(A)$, $C(A^T)$, $N(A^T)$.

Descomposición del valor singular (SVD) $A = U\Sigma V^T$ = (U ortogonal) multiplicada por (diagonal Σ) multiplicada por (V^T ortogonal V^T) Las primeras r columnas de U y V son bases ortonormales de $C(A)$ y $C(A^T)$, con $Av_i = \sigma_iu_i$ y valor singular $\sigma_i > 0$. Las últimas columnas de U y V son bases ortonormales de los espacios nulos de A^T y A.

Descomposición polar $A = QH$ Q ortogonal, H positiva (semi) definida.

Desigualdad de Schwarz $|v \cdot w| \leq \|v\|\,\|w\|$. Así, $|v^TAw|^2 \leq (v^TAv)(w^TAw)$ si $A = C^TC$.

Desigualdad del triángulo $\|u + v\| \leq \|u\| + \|v\|$. Para normas matriciales, $\|A + B\| \leq \|A\| + \|B\|$.

Desplazamiento cíclico S Permutación con $s_{21} = 1, s_{32} = 1, \ldots$, por último $s_{1n} = 1$. Sus valores característicos son n raíces $e^{2\pi ik/n}$ de 1; los vectores característicos son columnas de la matriz de Fourier F.

Determinante $|A| = \det(A)$ Definido por $\det I = 1$, con inversión de signo para intercambio de renglones, y linealidad en cada renglón. Así, $|A| = 0$ cuando A es singular. También, $|AB| = |A||B|$, $|A^{-1}| = 1/|A|$ y $|A^T| = |A|$. La gran fórmula para det (A) tiene una sumatoria de $n!$ términos, y en la fórmula de cofactores se utilizan determinantes de tamaño $n - 1$ y el volumen de la caja $= |\det(A)|$.

Diagonalización $\Lambda = S^{-1}AS\Lambda$ = Matriz de valores característicos y S = matriz de vectores característicos. A debe tener n vectores característicos independientes para que S sea invertible. Todas las $A^k = S\Lambda^kS^{-1}$.

Dimensión del espacio vectorial $\dim(\mathbf{V})$ = número de vectores que hay en cualquier base de $\mathbf{V}$.

Ecuación característica $\det(A - \lambda I) = 0$. Las n raíces son los valores característicos de A.

Ecuación normal $A^TA\hat{x} = A^Tb$ proporciona la solución por mínimos cuadrados de $Ax = 0$ si el rango total de A es n. La ecuación establece que (columnas de A) $\cdot$ $(b - A\hat{x}) = 0$.

Eigshow Valores característicos y valores singulares gráficos de 2 por 2. (MATLAB o Java).

Eliminación Sucesión de operaciones en los renglones que reduce A a una U triangular superior o a la forma reducida $R = \mathrm{rref}(A)$. Así, $A = LU$ con multiplicadores ℓ_{ij} en L, o $PA = LU$ con intercambio de renglones en P, o $EA = R$, con una E invertible.

Elipse (o elipsoide) $x^TAx = 1$ A debe ser positiva definida; los ejes de la elipse son vectores característicos de A, con longitudes $1/\sqrt{\lambda}$. (Para $\|x\| = 1$, los vectores $y = Ax$ están en la elipse $\|A^{-1}y\|^2 = y^T(AA^T)^{-1}y = 1$ mostrada por eigshow; las longitudes de los ejes son σ_i).

Espacio columna $C(A)$ Espacio de todas las combinaciones de las columnas de A.

Espacio nulo $N(A)$ Soluciones de $Ax = 0$. Dimensión $n - r$ = (# columnas) − rango.

Espacio nulo izquierdo $N(A^T)$ Espacio nulo de A^T = "espacio nulo izquierdo" de A porque $y^TA = 0^T$.

Espacio renglón $C(A^T)$ Todas las combinaciones de los renglones de A. Por convencionalismo, vectores columna.

Espacio vectorial V Conjunto de vectores tales que todas las combinaciones $cv + dw$ permanecen en $\mathbf{V}$. En la sección 2.1 se proporcionan ocho reglas necesarias para $cv + dw$.

Espectro de A Conjunto de valores característicos $\{\lambda_1, \ldots, \lambda_n\}$.

Exponencial La derivada de $e^{At} = I + At + (At)^2/2! + \cdots$ es Ae^{At}; $eA^t\,u(0)$ resuelve $u' = Au$.

Factorización $A = LU$ Si con la eliminación se llega de A a U *sin intercambio de renglones*, entonces la triangular inferior L con multiplicadores ℓ_{ij} (y $\ell_{ii} = 1$) regresa de U a A.

Factorización de Cholesky $A = CC^T = (L\sqrt{D})(L\sqrt{D})^T$ para A positiva definida.

Factorizaciones simétricas $A = LDL^T$ y $A = Q\Lambda Q^T$ El número de pivotes positivos en D y valores característicos positivos en Λ es el mismo.

Forma de Jordan $J = M^{-1}AM$ Si A tiene s vectores característicos independientes, su matriz "generalizada" M de vectores característicos proporciona $J = \mathrm{diag}(J_1, \ldots, J_s)$. El bloque J_k es $\lambda_kI_k + N_k$, donde N_k tiene unos en la diagonal 1. Cada bloque tiene un valor característico λ_k y un vector característico $(1, 0, \ldots, 0)$.

Forma escalonada reducida por renglones $R = \mathrm{rref}(A)$ Pivotes = 1; ceros arriba y abajo de los pivotes; r renglones diferentes de cero de R constituyen una base para el espacio renglón de A.

Gráfica de G Conjunto de n nodos unidos por parejas mediante m aristas. Una **gráfica completa** contiene todas las $n(n - 1)/2$ aristas entre los nodos. Un árbol sólo contiene $n - 1$ aristas y no contiene circuitos cerrados.

Gran fórmula para determinantes de n por n $\det(A)$ es una sumatoria de $n!$ términos, uno para cada permutación P de las columnas. Ese término es el producto $a_{1\alpha} \cdots a_{n\omega}$ en la diagonal de la matriz reordenada, multiplicado por $\det(P) = \pm 1$.

Inversa derecha A^+ Si el rango total del renglón de A es m, entonces $A^+ = A^T(AA^T)^{-1}$ tiene $AA^+ = I_m$.

Inversa izquierda A^+ Si el rango columna completo de A es n, entonces $A^+ = (A^TA)^{-1}A^T$ tiene $A^+A = I_n$.

La **multiplicación por bloques** de AB es permitida si las formas de los bloques lo permiten (las columnas de A y los renglones de B deben estar en bloques que correspondan).

Ley asociativa $(AB)C = A(BC)$ Los paréntesis pueden quitarse para dejar ABC.

Ley distributiva $A + (B + C) = AB + AC$ Se suma y luego se multiplica, o se multiplica y luego se suma.

Leyes de Kirchhoff *Ley de la corriente*: La corriente neta (en salida menos) es cero en cada nodo. *Ley del voltaje*: La suma de las diferencias de potencial (caídas de tensión) es cero en cualquier circuito cerrado.

Longitud $\|x\|$ Raíz cuadrada de x^Tx (teorema de Pitágoras en n dimensiones).

Matrices de conmutación $AB = BA$ En caso de ser diagonalizables, comparten n vectores característicos.

Matrices semejantes A y B $B = M^{-1}AM$ tiene los mismos valores característicos que A.

Matriz A positiva definida Matriz simétrica con valores característicos positivos y pivotes positivos. Definición: $x^TAx > 0$ a menos que $x = 0$.

Matriz A semidefinida (Positiva) semidefinida significa simétrica con $x^TAx \geq 0$ para todos los vectores x. Por tanto, todos los valores característicos $\lambda \geq 0$; no hay pivotes negativos.

Matriz $A = uv^T \neq 0$ con rango 1 Los espacios columna y renglón = rectas cu y cv.

Matriz acompañante En el renglón n se escribe $c_1, \ldots, c_n$ y en la diagonal se escriben $n - 1$ unos. Entonces $\det(A - \lambda I) = \pm(c_1 + c_2\lambda + c_3\lambda^2 + \cdots)$.

Matriz aleatoria rand(n) o randn(n) MATLAB crea una matriz con elementos aleatorios distribuido uniformemente en [0 1] para rand, y con una distribución normal estándar para randn.

Matriz aumentada $[A \ \ b]$ $Ax = b$ es resoluble cuando b está en el espacio columna de A; así, el rango de $[A \ \ b]$ es el mismo que el de A. La eliminación en $[A \ \ b]$ preserva correctas las ecuaciones.

Matriz circulante C Diagonales constantes se envuelven como en un desplazamiento cíclico. Toda matriz circulante es $c_0I + c_1S + \cdots + c_{n-1}S^{n-1}$. Cx = **convolución** $c * x$. Los vectores característicos están en F.

Matriz de adyacencia de una gráfica Matriz cuadrada con $a_{ij} = 1$ cuando hay un lado que va del nodo i al nodo j; en caso contrario, $a_{ij} = 0$. $A = A^T$ para una gráfica no dirigida.

Matriz de covarianza Σ Cuando las variables aleatorias x_i satisfacen media = valor medio = 0, sus covarianzas Σ_{ij} son los promedios de x_ix_j. Con las medias $\bar{x}_i$, la matriz Σ = media de $(x - \bar{x})(x - \bar{x})^T$ es positiva (semi)definida: es una matriz diagonal si las x_i son independientes.

Matriz de eliminación = Matriz elemental E_{ij} La matriz identidad con un $-\ell_{ij}$ en el elemento i, j ($i \neq j$). Así, E_{ij} resta ℓ_{ij} veces el renglón j de A del renglón i.

Matriz de Hilbert hilb(n) Elementos $H_{ij} = 1/(i + j - 1) = \int_0^1 x^{i-1}x^{j-1}dx$. Positiva definida pero $\lambda_{mín}$ extremadamente pequeño y número de condición grande.

Matriz de incidencia de una gráfica dirigida La matriz de incidencia de m por n aristas-nodos tiene un renglón por cada arista (del nodo i al nodo j), con elementos -1 y 1 en las columnas i y j.

Matriz de Pascal P_S = pascal(n) La matriz simétrica con elementos binomiales $\binom{i+j-2}{i-1}$. Todos los $P_S = P_LP_U$ contienen el triángulo de Pascal con det = 1 (consulte el índice para encontrar más propiedades).

Matriz de permutación P Hay $n!$ órdenes de $1, \ldots, n$; las $n!$ Ps tienen los renglones de I en el mismo orden. PA coloca las columnas de A en el mismo orden P es un producto de intercambios de renglones P_{ij}; P es *par* o *impar* (det $P = 1$ o -1) con base en el número de intercambios.

Matriz de rigidez K Cuando x proporciona los movimientos de los nodos en una estructura discreta, Kx proporciona las fuerzas internas. A menudo $K = A^{\mathrm{T}}CA$, donde C contiene constantes de resorte de la ley de Hooke y $Ax =$ alargamiento (esfuerzos) de los movimientos x.

Matriz de rotación $R = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$ hace rotar el plano un ángulo θ, y $R^{-1} = R^{\mathrm{T}}$ lo hace girar un ángulo $-\theta$. Matriz ortogonal, valores característicos $e^{i\theta}$ y $e^{-i\theta}$, vectores característicos $(1, \pm i)$.

Matriz diagonal D $d_{ij} = 0$ si $i \neq j$. **Diagonal por bloques**: ceros fuera de los bloques cuadrados D_{ii}.

Matriz diagonalizable A Debe tener n vectores característicos independientes (en las columnas de S; en forma automática con n valores característicos diferentes). Así, $S^{-1}AS = \Lambda =$ matriz de valores característicos.

Matriz en bloque Una matriz puede separarse en matrices por bloques, cortando entre renglones y/o entre columnas.

Matriz escalonada U El primer elemento diferente de cero (el pivote) en cada renglón aparece después del primer pivote en el renglón previo. Todos los renglones cero vienen al último.

Matriz F de Fourier Los elementos $F_{jk} = e^{2\pi ijk/n}$ proporcionan columnas ortogonales $\overline{F}^{\mathrm{T}}F = nI$. Por tanto, $y = Fc$ es la transformada discreta (inversa) de Fourier $y_j = \sum c_k e^{2\pi ijk/n}$.

Matriz H de Henkel Constante a lo largo de cada antidiagonal; h_{ij} depende de $i + j$.

Matriz H de Hessenberg Matriz triangular con una diagonal adyacente adicional diferente de cero.

Matriz hermitiana $A^{\mathrm{H}} = \overline{A}^{\mathrm{T}} = A$ Análogo complejo de una matriz simétrica $\overline{a_{ji}} = a_{ij}$.

Matriz hipercubo P_L^2 En el renglón $n + 1$ se cuentan las esquinas, las aristas, las caras, etc., de un cubo en $\mathbf{R}^n$.

Matriz identidad I (o I_n) Elementos diagonales $= 1$, elementos fuera de la diagonal $= 0$.

Matriz indefinida Matriz simétrica con valores característicos de ambos signos (+ y −).

Matriz inversa A^{-1} Matriz cuadrada con $A^{-1}A = I$ y $AA^{-1} = I$. No tiene inversa si det $A = 0$ y rango $(A) < n$; además, $Ax = 0$ para un vector x diferente de cero. Las inversas de AB y A^{T} son $B^{-1}A^{-1}$ y $(A^{-1})^{\mathrm{T}}$. La fórmula de cofactores es $(A^{-1})_{ij} = C_{ji}/\det A$.

Matriz M de cambio de base Los vectores v_j de la base anterior son combinaciones $\Sigma m_{ij} w_i$ de los vectores de la base nueva. Las coordenadas de $c_1v_1 + \cdots + c_nv_n = d_1w_1 + \cdots + d_nw_n$ están relacionadas según $d = Mc$. (Para $n = 2$, se hace $v_1 = m_{11}w_1 + m_{21}w_2$, $v_2 = m_{12}w_1 + m_{22}w_2$.)

Matriz M de Markov Todos los $m_{ij} \geq 0$ y la suma de cada columna es 1. El valor característico máximo es $\lambda = 1$. Si $m_{ij} > 0$, las columnas de M^k tienden al vector característico de estado estacionario $Ms = s > 0$.

Matriz normal N $NN^{\mathrm{T}} = N^{\mathrm{T}}N$, que conduce a vectores característicos (complejos) ortonormales.

Matriz N del espacio nulo Las columnas de N son las $n - r$ soluciones especiales de $As = 0$.

Matriz nilpotente N Alguna potencia de N es la matriz cero, $N^k = 0$. El único valor característico es $\lambda = 0$ (repetido n veces). Ejemplos: matrices triangulares con diagonal cero.

Matriz ortogonal Q Matriz ortogonal con columnas ortonormales, de modo que $Q^{\mathrm{T}}Q = I$ implica $Q^{\mathrm{T}} = Q^{-1}$. Preserva ángulos y longitud. $\|Qx\| = \|x\|$ y $(Qx)^{\mathrm{T}}(Qy) = x^{\mathrm{T}}y$ Todos los $|\lambda| = 1$, con vectores característicos ortogonales. Ejemplos: Rotación, reflexión, permutación.

Matriz proyección P sobre el subespacio S La proyección $p = Pb$ es el punto más próximo a b en S; el error $e = b - Pb$ es perpendicular a S. $P^2 = P = P^{\mathrm{T}}$, los valores característicos son 1 o 0; los vectores característicos están en S o en $\mathrm{S}^{\perp}$. Si las columnas de $A =$ base de S, entonces $P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$.

Matriz reflexión $Q = I - 2uu^{\mathrm{T}}$ El vector unitario u se refleja en $Qu = -u$. Todos los vectores x en el plano $u^{\mathrm{T}}x = 0$ permanecen sin cambio porque $Qx = x$. La "matriz del inquilino" cumple $Q^{\mathrm{T}} = Q^{-1} = Q$.

Matriz simétrica A La traspuesta es $A^{\mathrm{T}} = A$, y $a_{ij} = a_{ji}$. A^{-1} también es simétrica. Todas las matrices de la forma $R^{\mathrm{T}}R$, LDL^{T} y $Q\Lambda Q^{\mathrm{T}}$ son simétricas. Las matrices simétricas tienen valores característicos reales en Λ y vectores característicos ortonormales en Q.

Matriz simétrica sesgada K La traspuesta es $-K$, ya que $K_{ij} = -K_{ji}$. Los valores característicos son puramente imaginarios, los vectores característicos son ortogonales; e^{Kt} es una matriz ortogonal.

Matriz singular A Matriz cuadrada que no tiene inversa: det $(A) = 0$.

Matriz T de Toeplitz Matriz con diagonal constante, de modo que t_{ij} sólo depende de $j - i$. Las matrices de Toeplitz representan filtros lineales invariantes en el tiempo en el procesamiento de señales.

Matriz traspuesta A^{T} Los elementos $A^{\mathrm{T}}_{ij} = A_{ji}$. A^{T} es de n por m; $A^{\mathrm{T}}A$ es cuadrada, simétrica y positiva semidefinida. Las traspuestas de AB y A^{-1} son $B^{\mathrm{T}}A^{\mathrm{T}}$ y $(A^{\mathrm{T}})^{-1}$.

Matriz tridiagonal T $t_{ij} = 0$ si $|i - j| > 1$. T^{-1} tiene rango 1 arriba y abajo de la diagonal.

Matriz unitaria $U^{\mathrm{H}} = \overline{U}^{\mathrm{T}} = U^{-1}$ Columnas ortonormales (análogo complejo de Q).

Matriz V de Vandermonde $Vc = b$ proporciona el polinomio $p(x) = c_0 + \cdots + c_{n-1}x^{n-1}$ con $p(x_i) = b_i$ en n puntos. $V_{ij} = (x_i)^{j-1}$, y det $V =$ producto de $(x_k - x_i)$ para $k > i$.

Método de Gauss-Jordan A se invierte con operaciones en los renglones sobre $[A \;\; I]$ para llegar a $[I \;\; A^{-1}]$.

Método simplex de programación lineal El vector x^* de costo mínimo se encuentra desplazándose de un vértice hacia el vértice de menor costo a lo largo de las aristas del conjunto posible (donde se satisfacen las restricciones $Ax = b$ y $x \geq 0$). ¡El costo mínimo se encuentra en un vértice!

Método iterativo Secuencia de pasos que se siguen para aproximarse a la solución deseada.

Método del gradiente conjugado Sucesión de pasos para resolver una positiva definida $Ax = b$ minimizando $\frac{1}{2}x^{\mathrm{T}}Ax - x^{\mathrm{T}}b$ sobre subespacios crecientes de Krylov.

Multiplicación $Ax = x_1$(columna 1) $+ \cdots + x_n$(columna n) $=$ combinación de columnas.

Multiplicación de matrices AB El elemento i, j de AB es (renglón i de A) · (columna j de B) $= \Sigma a_{ik}b_{kj}$. Por columnas: columna j de $AB = A$ multiplicada por la columna j de B. Por renglones: el renglón i de A multiplica a B. Columnas por renglones: $AB =$ suma de (columna k)(renglón k). Todas estas definiciones equivalentes provienen de la regla de que AB multiplicada por x es igual a A multiplicada por Bx.

Multiplicador ℓ_{ij} El renglón pivote j se multiplica por ℓ_{ij} y se resta del renglón i para eliminar el elemento i, j: $\ell_{ij} =$ (elemento a eliminar)/(j-ésimo pivote).

Multiplicidades AM y GM La multiplicidad algebraica AM de un valor característico λ es el número de veces que λ aparece como raíz de det $(A - \lambda I) = 0$. La multiplicidad geométrica GM es el número de vectores independientes (= dimensión del espacio característico para λ).

Norma $\|A\|$ de una matriz La "norma ℓ^2" es la razón máxima $\|Ax\|/\|x\| = \sigma_{\text{máx}}$. Así, $\|Ax\| \leq \|A\|\|x\|$, $\|AB\| \leq \|A\|\|B\|$, y $\|A + B\| \leq \|A\| + \|B\|$. **Norma de Frobenius** Las normas $\|A\|_F^2 = \sum\sum a_{ij}^2$; ℓ^1 y ℓ^∞ son las máximas sumas de columnas y renglones de $|a_{ij}|$.

Número de condición $\operatorname{cond}(A) = \kappa(A) = \|A\|\|A^{-1}\| = \sigma_{\text{máx}}/\sigma_{\text{mín}}$. En $Ax = b$, el cambio relativo $\|\delta x\|/\|x\|$ es menos que cond(A) multiplicado por el cambio relativo $\|\delta b\|/\|b\|$. Los números de condición miden la *sensibilidad* de la salida al cambio en la entrada.

Números de Fibonacci $0, 1, 1, 2, 3, 5, \ldots$, que satisfacen $F_n = F_{n-1} + F_{n-2} = (\lambda_1^n - \lambda_2^n)/(\lambda_1 - \lambda_2)$. La razón de crecimiento $\lambda_1 = \left(1 + \sqrt{5}\right)/2$ es el valor característico máximo de la matriz de Fibonacci $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.

Números de Lucas $L_n = 2, 1, 3, 4, \ldots$, satisfacen $L_n = L_{n-1} + L_{n-2} = \lambda_1^n + \lambda_2^n$ con valores característicos $\lambda_1, \lambda_2 = (1 \pm \sqrt{5})/2$ de la matriz de Fibonacci $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.

Ondas $w_{jk}(t)$ o vectores w_{jk} Reescalan y desplazan el eje del tiempo para crear $w_{jk}(t) = w_{00}(2^jt - k)$. Los vectores provenientes de $w_{00} = (1, 1, -1, -1)$ deben ser $(1, -1, 0, 0)$ y $(0, 0, 1, -1)$.

Ortogonalización de Gram-Schmidt $A = QR$ Columnas independientes en A, columnas ortonormales en Q. Cada columna q_j de Q es una combinación de las j primeras columnas de A (y al revés, de modo que R es triangular superior). La simbología es diag $(R) > 0$.

Pivote d El primer elemento diferente de cero cuando un renglón se utiliza en eliminación.

Plano (o hiperplano) en $\mathbf{R}^n$ Las soluciones de $a^{\mathrm{T}}x = 0$ proporcionan el plano (dimensión $n - 1$) perpendicular a $a \neq 0$.

Pivoteo parcial Durante la eliminación, el j-ésimo pivote se elige como el elemento más grande disponible (en valor absoluto) en la columna j. Así, todos los multiplicadores satisfacen $|\ell_{ij}| \Sigma 1$. El error por redondeo está controlado (dependiendo del *número de condición* de A).

Polinomio mínimo de A El polinomio de menor grado con $m(A) =$ matriz cero. Las raíces de m son valores característicos, y $m(\lambda)$ divide a det $(A - \lambda I)$.

Producto cruz $u \times v$ en $\mathbf{R}^3$ Vector perpendicular a u y v, de longitud $\|u\|\|v\||\operatorname{sen}\theta| =$ área del paralelogramo, calculada como el "determinante" de $[i \;\; j \;\; k;\; u_1 \;\; u_2 \;\; u_3;\; v_1 \;\; v_2 \;\; v_3]$.

Producto de Kronecker (producto tensorial) $A \otimes B$ Bloques $a_{ij}B$, valores característicos $\lambda_p(A)\lambda_q(B)$.

Producto punto $x^{\mathrm{T}}y = x_1y_1 + \cdots + x_ny_n$ El producto punto complejo es $\overline{x}^{\mathrm{T}}y$. El producto punto de vectores perpendiculares es cero. $(AB)_{ij} =$ (renglón i de A)·(columna j de B).

Producto externo uv^{T} Columna multiplicada por renglón $=$ matriz de rango 1.

Proyección $p = a(a^{\mathrm{T}}b/a^{\mathrm{T}}a)$ sobre la recta que pasa por a El rango de $P = aa^{\mathrm{T}}/a^{\mathrm{T}}a$ es 1.

Punto silla de $f(x_1, \ldots, x_n)$ Punto en que las primeras derivadas de f son cero y la segunda matriz de derivadas ($\partial^2 f/\partial x_i \partial x_j =$ **matriz hessiana**) es indefinida.

Radio espectral $= |\lambda_{\text{máx}}|$.

Rango $A(r)$ Es igual al número de pivotes $=$ dimensión del espacio columna $=$ dimensión del espacio renglón.

Rango total de la columna $r = n$ Columnas independientes, $N(A) = \{0\}$, sin variables libres.

Rango total del renglón $r = m$ Renglones independientes, por lo menos una solución de $Ax = b$; el espacio columna es todo $\mathbf{R}^m$. Rango total significa **rango total de la columna** o **rango total del renglón**.

Red Gráfica dirigida que tiene $c_1, \ldots, c_m$ constantes asociadas con las $\|A\|$ aristas.

Regla de Cramer para $Ax = b$ B_j tiene b reemplazando la columna j de A y $x_j = |B_j|/|A|$.

Representación por columna de $Ax = b$ El vector b se convierte en una combinación de las columnas de A. El sistema sólo es resoluble cuando b está en el espacio columna $C(A)$.

Representación por renglón de $Ax = b$ Cada ecuación proporciona un plano en $\mathbf{R}^n$; los planos se cortan en x.

Seudoinversa A^+ (inversa de Moore-Penrose). La matriz n por m que "invierte" a A del espacio columna al espacio renglón, con $N(A^+) = N(A^T)$. A^+A y AA^+ son las matrices proyección sobre el espacio renglón y el espacio columna. Rango (A^+) = rango (A).

Sistema resoluble $Ax = b$ El miembro derecho b está en el espacio columna de A.

Solución completa $x = x_p + x_n$ de $Ax = b$ (x_p particular) + (x_n en el espacio nulo).

Solución particular x_p Cualquier solución de $Ax = b$; a menudo x_p tiene variables libres = 0.

Solución por mínimos cuadrados $\widehat{x}$ El vector $\widehat{x}$ que minimiza el error $\|e\|^2$ resuelve $A^TA\widehat{x} = A^Tb$. Así, $e = b - A\widehat{x}$ es ortogonal a todas las columnas de A.

Soluciones especiales de $As = 0$. Una variable libre es $s_i = 1$, las demás variables libres = 0.

Subespacio de Krylov $K_j(A, b)$ Subespacio generado por $b, Ab, \ldots, A^{j-1}b$. Métodos numéricos aproximan $A^{-1}b$ por x_j con residuos $b - Ax_j$ en este subespacio. Una buena base de K_j requiere sólo multiplicación por A en cada paso.

Subespacios ortogonales Todo v en V es ortogonal a todo w en W.

Subespacio S de V Cualquier espacio vectorial dentro de $\mathbf{V}$, incluyendo $\mathbf{V}$ y $\mathbf{Z}$ = {vector cero}.

Suma $\mathbf{V} + \mathbf{W}$ de subespacios Espacio de todos (los v en $\mathbf{V}$) + (los w en $\mathbf{W}$). **Suma directa**: $\dim(\mathbf{V} + \mathbf{W}) = \dim \mathbf{V} + \dim \mathbf{W}$, cuando $\mathbf{V}$ y $\mathbf{W}$ sólo comparten el vector cero.

Sustitución hacia atrás Los sistemas triangulares superiores se resuelven en orden inverso, de x_n a x_1.

Suma vectorial $v + w = (v_1 + w_1 + \ldots + v_n + w_n)$ = diagonal del paralelogramo.

Teorema de Cayley-Hamilton $p(\lambda) = \det(A - \lambda I)$ tiene $p(A)$ = *matriz cero*.

Teorema fundamental El espacio nulo $N(A)$ y el espacio renglón $C(A^T)$ son complementos ortogonales (subespacios perpendiculares de $\mathbf{R}^n$ con dimensiones r y $n - r$) de $Ax = 0$. Aplicado a A^T, el espacio columna $C(A)$ es el complemento ortogonal de $N(A^T)$.

Teorema espectral $A = Q\Lambda Q^T$ A simétrica real tiene λ_i real y q_i ortonormal, con $Aq_i = \lambda_i q_i$. En mecánica, q_i proporciona los *ejes principales*.

Transformada de Fourier rápida (TFR) Factorización de la matriz de Fourier F_n en matrices S_i $\ell = \log_2 n$ multiplicadas por una permutación. Cada S_i sólo requiere $n/2$ multiplicaciones, de modo que $F_n x$ y $F_n^{-1}c$ pueden calcularse con $n\ell/2$ multiplicaciones. Es revolucionaria.

Transformación afín $T(v) = Av + v_0$ = transformación lineal más desplazamiento.

Transformación lineal T Cada vector v en el espacio de entrada se transforma en $T(v)$ en el espacio de salida, y la linealidad requiere $T(cv + dw) = cT(v) + dT(w)$. Ejemplos: multiplicación matricial Av, diferenciación en el espacio de funciones.

Traza de A Suma de los elementos en la diagonal = suma de los valores característicos de A. Tr AB = Tr BA.

Una **gráfica dirigida** tiene una flecha especificada en cada arista.

$v_1, \ldots, v_n$ linealmente dependientes Una combinación diferente de todos los $c_i = 0$ proporciona $\Sigma c_i v_i = 0$.

Valor característico λ y vector característico x $Ax = \lambda x$ con $x \neq 0$, de modo que $\det(A - \lambda I) = 0$.

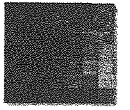
Variable libre x_i La columna i carece de pivote en la eliminación. Es posible asignar cualesquiera valores a las $n - r$ variables libres, y luego $Ax = b$ determina las r variables pivote (¡en caso de ser resoluble!)

Vector v en $\mathbf{R}^n$ Sucesión de n números reales $v = (v_1, \ldots, v_n)$ = punto en $\mathbf{R}^n$.

Vectores independientes $v_1, \ldots, v_k$ Ninguna combinación $c_1v_1 + \cdots + c_kv_k$ = vector cero a menos que todas las $c_i = 0$. Si las vs son las columnas de A, la única solución de $Ax = 0$ es $x = 0$.

Vectores ortonormales $q_1, \ldots, q_n$ Los productos punto son $q_i^Tq_j = 0$, si $i \neq j$ y $q_i^Tq_i = 1$. La matriz Q con estas columnas ortonormales cumple $Q^TQ = I$. Si $m = n$, entonces $Q^T = Q^{-1}$ y $q_1, \ldots, q_n$ es una base ortonormal de $\mathbf{R}^n$; toda $v = \sum(v^Tq_j)q_j$.

Volumen de una caja Los renglones (o las columnas) de A generan una caja con volumen $|\det(A)|$.



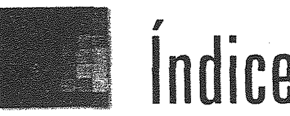
Códigos de enseñanza MATLAB

cofactor	Calcula la matriz de cofactores de n por n.
cramer	Resuelve el sistema $Ax = b$ con la regla de Cramer.
deter	Matriz de determinantes calculada a partir de los pivotes en $PA = LU$.
eigen2	Valores característicos, vectores característicos y det $(A - \lambda I)$ para matrices de 2 por 2.
eigshow	Demostración gráfica de valores característicos y valores singulares.
eigval	Valores característicos y su multiplicidad como raíces de det $(A - \lambda I) = 0$.
eigvec	Calcula tantos vectores característicos linealmente independientes como es posible.
elim	Reducción de A a forma escalonada reducida R por medio de una E invertible.
findpiv	Encuentra un pivote para eliminación gaussiana (utilizado por **plu**).
fourbase	Construye bases para todos los cuatro subespacios fundamentales.
grams	Ortogonalización mediante el proceso de Gram-Schmidt de las columnas de A.
house	Matriz de 2 por 12 que proporciona las coordenadas de los vértices de una casa.
inversa	Matriz inversa (en caso de existir) por eliminación gaussiana.
leftnull	Calcula una base para el espacio nulo izquierdo.
linefit	Grafica el ajuste por mínimos cuadrados a m puntos dados por medio de una recta.
lsq	Solución por mínimos cuadrados de $Ax = b$ a partir de $A^TA\widehat{x} = A^Tb$.
normal	Valores característicos y vectores característicos ortonormales cuando $A^TA = AA^T$.
nulbasis	Matriz de soluciones especiales de $Ax = 0$ (base para el espacio nulo).
orthcomp	Encuentra una base para el complemento ortogonal de un subespacio.
partic	Solución particular de $Ax = b$, con las tres variables libres iguales a cero.
plot2d	Gráfica bidimensional para las figuras de las casas.
plu	Factorización rectangular $PA = LU$ con intercambios de renglones.
poly2str	Expresa un polinomio como una cadena.
project	Proyecta un vector b sobre el espacio columna de A.
projmat	Construye la matriz proyección sobre el espacio columna de A.
randperm	Construye una permutación aleatoria.
rowbasis	Calcula una base para el espacio renglón a partir de los renglones pivote de R.
samespan	Prueba si dos matrices tienen el mismo espacio columna.
signperm	Determinante de la matriz permutación con renglones ordenados por p.
slu	Factorización LU de una matriz cuadrada sin usar *ningún intercambio de renglones*.
slv	Aplica **slu** para resolver el sistema $Ax = b$ sin permitir intercambios de renglones.
splu	Factorización cuadrada $PA = LU$ *usando intercambios de renglones*.
splv	Solución de un sistema cuadrado invertible $Ax = b$.
symmeig	Calcula los valores característicos y los vectores característicos de una matriz simétrica.
tridiag	Construye una matriz tridiagonal con diagonales constantes a, b, c.

Estos códigos de enseñanza están disponibles directamente a partir de la Linear Algebra Home Page:

`http://web.mit.edu/18.06/www.`

Fueron escritos en MATLAB y traducidos a Maple y Mathematica.



Índice

$A = LDL^{\mathrm{T}}$, 51, 60, 319-320, 325, 474, 480
$A = LDU$, 36, 51, 224, 369, 474
$A = LU$, 34-35
$A = MJM^{-1}$, 300, 474
$A = Q\Lambda Q^{\mathrm{T}}$, 285, 288, 297-298, 320-323, 474, 480
$A = QR$, 174, 179, 181-182, 351, 363, 474, 477
$A = QS$, 333
$A = U\Sigma V^{\mathrm{T}}$, 306, 331-333, 336, 474, 480
$A = S\Lambda S^{-1}$, 245, 250, 255, 257, 267, 300, 474
AA^{T}, 46, 108, 162, 222-223, 306, 331-336, 357, 475
$A^{\mathrm{T}}A$, 45, 108-109, 114, 161-168, 179, 182, 184, 306, 331-335, 341, 356-357, 363, 475, 481, 488
$A^{\mathrm{T}}CA$, 120-124, 480
$\mathbf{C}^n$, 248, 273, 280, 282, 288, 292
e^{At}, 266-279
$PA = LU$, 38-39
$Q\Lambda Q^{\mathrm{T}}$, 320-323, 327
RR^{T} y $R^{\mathrm{T}}R$, 51-52
$\mathbf{R}^n$, 69, 72-73, 288
$S^{-1}AS$, 132, 245-248, 285, 293, 299, 301, 324, 477

A
$A = LU$, 34, 35
Abel, Niels Henrik, 239
Álgebra booleana, 204
Algoritmo avaricioso, 405
Algoritmo Crout, 36
Algoritmo QR, 351, 359, 364-365
Análisis de regresión, 153
Análisis de sensibilidad, 396
Ancho de banda, 61, 371-372
Apareamiento completo, 403
Apareamiento, 403-407, 472, 476
Aplicaciones de los determinantes, 201, 220-229
Applied Mathematics and Scientific Computing, 122, 320-321, 349
Árbol, 117, 123-124, 255, 405, 407
Árbol generador, 117
Árbol generador más corto, 405
Área, 137, 223-229, 349, 454-455, 477
Arnoldi, 374

B
Base estándar, 174
Base, 95, 141
Bringing Down the House, 377, 412
Buniakowsky, 155

C
Cadena de vectores característicos, 423, 427
Cálculo de A^{-1}, 46-47
California, 257-258, 381
Cambio de base, 132, 136, 294-295, 302, 476
Cambio de variables, 293, 390, 426
Capacidad, 119, 401-406, 472
Capacidad de la arista, 119
Casos singulares, 3, 7-11, 13
Cauchy-Buniakowsky-Schwarz, 155
Cayley-Hamilton, 253, 304, 427, 456-457, 476
$CD = -DC$, 27, 206, 231, 302
Cero en una posición pivote, 13, 28, 33, 37-38, 42, 48-49, 78-84, 89, 105, 202, 474
Ciclo, 114-124, 146, 374, 405, 444, 477-478
Circunferencia unitaria, 190, 282, 298
$\mathbf{C}^n$, 248, 280, 282, 288, 292
Cofactores, 213
Columna a la vez, 21, 26, 46, 129, 331, 423
Columnas multiplicadas por renglones, 30, 285, 333, 478
Combinación de columnas, 6-7, 71-72, 92, 478
Combinación de renglones, 429
Combinación lineal, 6-7
Complemento de Schur, 31, 219, 431, 475, 480
Complemento, 145-152
Completamiento al cuadrado, 313, 316-317, 345
Composición, 131
Composición continua, 254
Condición a la frontera, 59, 64, 347, 350
Condición de Hall, 404
Condición natural a la frontera, 59, 64, 347, 350
Condiciones de Kuhn-Tucker, 394, 397
Condiciones flojas complementarias, 394, 409
Conductancia, 119
Congruencia, 324, 326
Conjugados complejos, 281
Conjunto factible, 378, 382
Conjunto generador minimal, 97
Conjunto independiente maximal, 97
Conmutativa, 23, 25, 69
Cono, 399-400
Constantes arbitrarias, 59, 115
Convergencia, 368
Cooley, 194
Coordenada(o), 6, 69-70, 201, 229, 282
Coordenadas polares, 282, 289, 333
Corte mínimo, 402
Coseno, 102, 152-159, 182-184, 188-191, 198, 272, 274
Costo de la eliminación, 14, 15
Costos reducidos, 386, 396

D
Dantzig, George Bernard, 382
De dimensión infinita, 69, 347
Defectuoso(a), 268, 293, 299
Degeneración, 385, 395
Dependiente, 9-11, 80-82, 92-111, 116-117, 259, 282, 333-335
Descenso más pronunciado, 390
Descomposición de Cholesky, 320
Descomposición del valor singular (DVS), 331-337
Descomposición, 32, 148, 298, 331-338, 357, 363, 475, 479-480
Desigualdad de Schwarz, 154-155, 183, 250
Desigualdad del triángulo, 157, 262, 358, 480
Desigualdades, 377-381
Determinante cero, 204
Determinante jacobiano, 201
Determinantes
fórmulas, 201, 210-219
propiedades, 203-209
"razón de determinantes", 1, 202, 224
Diagonalizable, 238, 246, 249-253, 270, 290, 296-303, 306-308, 427, 457, 473, 476-477
Diagonalización de matrices, 245
forma de Jordan, 422-427
simultánea, 326
transformaciones de semejanza, 301
Diagonalización simultánea, 326
Diagonalmente dominante, 373
Diferencias finitas, 61, 64, 270, 346, 348, 354, 370, 418
Difusión, 268
Dilema del prisionero, 412
Dimensión, 69-73, 81-96, 104-106, 147, 181-183, 314-315, 416
del espacio columna, 98
del subespacio, 81
del espacio vectorial, 96
Dirección de la dimensionalidad, 371
Distancia, 152, 155-157, 161, 165-166, 173
Distancias verticales, 166
Dualidad débil, 393
DVS. *Véase* descomposición del valor singular

E
e^{At}, 266-279
Economía, 58, 153, 260-263, 265, 379, 396, 399
Ecuación de Bellman, 406
Ecuación de Laplace, 418
Ecuación de onda, 275
Ecuación del calor, 270
Ecuación en diferencias, 59, 64, 193, 238, 250, 254-270, 273-275, 293, 348, 359, 367, 419
Ecuación homogénea, 73
Ecuación normal, 162
Ecuaciones diferenciales
análisis de Fourier, 122
cambio a ecuaciones matriciales, 59
difusión, 268 y e^{At}, 266-279
ecuaciones de segundo orden, 274
ecuación diferencial parcial de Laplace, 418
estabilidad, 270, 273
inestabilidad, 270, 271, 273
superposición, 237
transformaciones de semejanza, 293
Ecuaciones diferenciales parciales, 371, 418
eigshow, 240
Einstein, Albert, 21
Ejes derechos, 175, 223
Ejes principales, 334
Eliminación gaussiana, 1-68
$A = LU$ y $PA = LU$, 34-35, 38-39
casos singulares, 7-11, 13
geometría de ecuaciones lineales, 3-10
notación matricial, 19-31
ortogonalidad, 160, 184
Eliminación hacia delante, 32, 36
Eliminación por bloques, 120, 219, 480
Eliminación, 1, 9
Elipses y elipsoides
eigshow, 240
espacio de Hilbert, 182, 183
método de Khachian, 389
matrices positivas definidas, 322
teorema del eje principal, 285
Energía, 272-275, 287, 334, 339-340, 347-350
Entrelazamiento, 343-344
Equilibrio, 120, 122, 261, 344, 472
Errores por redondeo, 61-63, 333, 352, 355-356, 359, 479
Escalar, 6, 19-71, 73-75, 126, 143, 234, 278, 282, 339, 415, 478
Esfuerzo cortante, 132-133
Espacio columna, 71, 72, 104, 107
Espacio de Hilbert, 182-183
Espacio euclidiano, 183
Espacio. *Véase* Espacio vectorial
Espacio nulo, 71, 73, 107, 144
Espacio nulo izquierdo, 107
Espacio renglón, 102-110, 116-117, 144-148, 331
Espacios de funciones, 183
Espacios vectoriales, 69-140
ortogonalidad, 141
producto, suma e intersección, 415-421
subespacios, 102-113
subespacios fundamentales, 102-113
transformación lineal, 125-137
Esquema de cinco puntos, 371
Estabilidad, 270, 273
Estadística, 122, 153, 162, 172, 325
Estado estacionario, 257-259, 261, 263-264, 273, 275, 306, 309, 360, 478
Estrategia mezclada, 408
Existencia y unicidad, 69
Existencia, 61, 69, 107-109, 410
Experimento, 19, 67, 153, 165-167
Exponencial pura, 426
Exponenciales, 266-279

F
Factorización, 36, 213
de Gram-Schmidt, 363
factor de sobrerrelajamiento, 369
L y U, 3
matriz de Fourier, 474
polar, 333
simétrica, 51
triangular, 32-44
Factorización LDL^{T}, 51-53, 60
Factorización LDU, 36-37, 41-43, 51-53, 60-63
Factorización LU, 33-44
Factorización polar, 333
Factorización QR. *Véase* Proceso de Gram-Schmidt
Falla, 7, 13, 16, 18, 49
Filippov, A. F., 423
Filtración, 189
Fix, George, 349
Forma de Jordan, 300, 422-427
Forma R reducida por renglones, 77-78
Forma U escalonada, 77-78
Fórmula de Euler, 117, 191
Fórmulas para determinantes, 201, 210-219
de Euler, 117, 191
pitagóricas, 142
producto de pivotes, 47, 202
Fredholm, 149
Freund, Robert, 398

Frobenius, 261-262, 479
Función característica, 270, 346, 349
Función de costo, 378
Fútbol, 118-119, 124, 322

G
Galois, Évariste, 239
Gauss-Jordan, 47-49
Generación de un espacio, 94
Geometría de planos, 2
Gershgorin, 373-374
Girasol, 255
Givens, 302
Golub, Gene Howard, 372
Gráfica dirigida, 104, 114
Gráficas y redes, 114-124, 401-407
Grupo, 58, 66-67, 80, 213, 330, 351, 402, 436, 465

H
Hiperplano de separación, 398-399
Homogénea, 20, 92, 149, 237, 439, 447

I
IBM, 15
Inconsistente, 8
Indefinida, 312-314, 322-323, 327-330, 464, 478, 480
Independencia lineal, 82-102
Independencia, 92-105, 143, 164, 330, 425
Inercia, ley de la, 324
Inestabilidad
ecuaciones en diferencias, 270, 271, 273
ecuación de Fibonacci, 259
errores por redondeo, 63
valores característicos y vectores característicos, 234, 259, 207, 273
Infinidad de soluciones, 3, 8, 9
Integración, 127, 183
Intercambios de renglones, 32-44
Interés compuesto, 254, 259
Intersección de espacios, 415-421
Introduction to Applied Mathematics, 122, 320, 349
Invariante, 324, 480
Inversa derecha, 338, 466
Inversa izquierda, 45, 177
Inversa, 45-48
de un producto, 34
de una traspuesta, 38, 45-58
fórmula para A^{-1}, 52, 221
Invertible = No singular, 48, 49

J
Juego de dos personas, 408
Juego de suma cero, 409

K
Kernel, 104, 135, 445

L
Lanczoz, 374-375
LAPACK (Linear Algebra PACKage), 351
Las Vegas, 377, 412
Lema de Schur, 296
Leontief, 260
Ley asociativa, 23, 29, 34, 46, 134, 445, 476
Ley de la corriente de Kirchhoff, 106, 116, 117, 120, 402
Ley de la corriente, 106, 116-117, 120-122, 401-402, 478
Ley de la inercia de Sylvester, 324
Ley de la inercia, 324
Ley de los cosenos, 152-159
Ley de Newton, 273
Ley de Ohm, 118-122
Ley de Pitágoras, 141, 154, 177, 335
Ley del voltaje de Kirchhoff, 115, 120, 146
Ley del voltaje. *Véase* Kirchhoff
Ley distributiva, 445, 477
Linealmente dependiente, 92
Longitud, 119, 404
LU incompleta, 372
Lyapunov, Aleksandr, 272

M
Mal acondicionado, 62-64, 184, 352-353, 436
Mantequilla de cacahuate, 380, 392-393
Mathematics for the Millions, 222
MATLAB, 211, 239, 285
Matrices complejas, 280-292
Matrices de incidencia arista-nodo, 104
Matrices de incidencia, 104, 118, 401
Matrices hermitianas sesgadas, 288, 298
Matrices no diagonalizables, 238, 246, 268, 293, 299
Matrices simétricas sesgadas, 410
Matriz (matrices)
acompañante, 476
banda, 59, 61, 401
circulante, 189
consumo, 260
de adyacencia, 124, 476
de coeficientes, 3, 5, 19-22, 59-60
de cofactores, 213-222
de covarianza, 169-172
de diferencias finitas, 61, 64, 270, 346-348, 370, 418
de Fourier, 176, 182-184, 188-195, 287, 419, 477
de Hilbert, 184
de incidencia, 104, 118, 401
de Jordan, 300-422
de Markov, 258, 261, 273, 360
de pagos, 408-413
de rango uno, 109-110, 156, 306, 479
de tablero de ajedrez, 139, 216, 242
de transición, 258
defectuosa, 238, 246, 268, 293, 299
de Hessenberg, 361, 365
diagonal, 36
diagonalizable, 246, 249
diferencia, 59, 115-119, 221
elemental, 22, 32, 49
eliminación de, 22, 32
escalonada, 77
exponencial, 234-237, 256, 266-274, 301, 306, 477
hermitiana, 280
hermitiana sesgada, 288, 298
identidad, 22
indefinida, 312-314, 327-330
inversa, 45-48
invertible, 48-49
mal acondicionada, 62-64, 184, 352-353, 436
multiplicación de, 19-31
nilpotente, 309, 479
no diagonalizable, 238, 246, 268, 293, 299
no negativa, 257-262, 378-382, 398-399
no singular, 9, 13
norma y número de condición, 352-358
normal, 162-170
notación, 2-3, 9, 19
ortogonal, 175
permutación, 203, 224, 403
positiva definida, 311-330
positiva, 60, 261
producto cruz, 177
proyección, 25, 164, 238
raíz cuadrada, 142, 181, 189-193, 223
rectangular, 20, 109, 114, 129, 177
reflexión, 125
rotación, 125, 131, 247, 365
semejante, 293
semidefinida, 314, 321-322, 333, 480
simétrica sesgada, 410
simétrica, 50-58
singular, 38, 204
traspuesta, 3, 45-51
triangular inferior, 33, 71
triangular, 35-36
tridiagonal, 60
unitaria, 286, 298, 331
Matriz acompañante, 242, 456, 476
Matriz alargamiento, 125
Matriz banda, 59, 61, 401
Matriz bidiagonal, 61, 363, 364
Matriz cero, 300
Matriz circulante, 189, 197, 291-292, 476
Matriz conectividad, 115
Matriz consumo, 260
Matriz de cinco puntos, 372
Matriz de cofactores, 218, 221, 226, 454
Matriz de covarianza, 169-172, 449, 476
Matriz de eliminación, 2, 22
Matriz de Fourier, 188, 190-192, 195-197, 287, 291, 309, 419, 421, 461, 475
Matriz de Hessenberg, 361, 365
Matriz de Hilbert, 184
Matriz de Householder, 361-365
Matriz de Markov, 244, 257-258, 261, 273, 360, 478
Matriz de pagos, 408-409, 413
Matriz de transición, 258-259, 263-264, 458
Matriz de Vandermonde, 109
Matriz de vectores característicos, 245, 247, 249, 251, 253, 291-293, 296, 331-332, 419, 477-478
Matriz defectuosa, 238, 246, 268, 293, 299
Matriz diagonal, 36, 46, 204-206, 238, 245, 267, 322, 327-335, 390, 415, 422
Matriz diferenciación, 128-129
Matriz elemental, 22, 32, 49
Matriz entrada-salida, 260
Matriz estable, 290, 332
Matriz hermitiana
característica de ser positiva definida, 334
valores característicos y vectores característicos, 280, 283-286, 288, 297, 298
Matriz identidad, 22
Matriz integración, 129
Matriz inversa, 45-48
Matriz masa, 321-325, 350, 406
Matriz nilpotente, 309, 479
Matriz no singular, 9, 13
Matriz normal, 298, 303, 357, 479
Matriz permutación, 203, 224, 403
Matriz positiva definida, 311-350
mínimos, 311-317
principios mínimos, 339-345
pruebas para la característica de ser positiva definida, 318-330
semidefinida, 314, 321
Matriz positiva, 60, 261
Matriz producto cruz, 177
Matriz proyección, 125
Matriz raíz cuadrada, 193, 320, 332, 334, 336
Matriz rala, 59, 348
Matriz reflexión, 125
Matriz rigidez, 119, 348
Matriz rotación, 125, 131, 247, 365
Matriz semejante, 294, 296, 306, 324, 361, 422, 480
Matriz simétrica, 50-58
$Q\Lambda Q^T$, 320-323, 327
simétrica LDL^T, 51
valores característicos y vectores característicos, 280, 286, 298
Matriz singular, 38, 204
Matriz topología, 115
Matriz traspuesta, 3, 45-51
Matriz triangular inferior, 33, 71
Matriz triangular superior, 32, 181
Matriz unitaria, 286, 298, 331
Matriz valor característico, 245, 247, 251, 292, 325, 331, 474-475, 477
Maximización del mínimo, 410
Mecánica cuántica, 249
Media, 178, 179
Media aritmética, 154, 447
Menores, 213
Método de Gauss-Seidel, 368-371
Método de Karmarkar, 390
Método de Khachian, 389
Método de potencias inverso desplazado, 360
método de potencias inverso Jacobi, 361, 368, 369, 371
Método de potencias inverso, 360
Método del elemento finito, 321, 346
Método del gradiente conjugado, 372, 390
Método del punto interior, 377, 390
Método iterativo, 367-372
Método simplex, 377, 379, 382-391
Método simplex revisado, 389
Mínimo global, 312
Mínimo local, 312
Mínimos cuadrados no lineales, 168
Mínimos cuadrados, 119, 153, 160-173, 177
Mínimos y máximos, 311-317
MIT, 118-119
Modelo de von Neumann, 262
Modo normal, 238, 275, 330
Multiplicación de matrices, 20-21, 131
Multiplicación por bloques, 224
Multiplicadores de Lagrange, 340, 396
Multiplicidad, 246, 301, 478, 481
Muro frontal, 144, 382
Mutuamente ortogonales, 143

N
Negativa definida, 314
Neutralmente estable, 259
New York Times, 119
No hay solución, 2, 3, 7, 8,
No negatividad, 378, 383, 398
Nodos, 104, 114-117
Norma de una matriz, 352
Notación sigma, 21
Núcleo, 104, 135, 445
Nulidad, 104-106, 127, 416
Número de pasos de eliminación, 3, 4, 239
Número de vectores en la base, 96
Números complejos, 189
Números de condición, 332
Números de Fibonacci, 238, 255, 256, 259

O

Operaciones aritméticas, 14, 15
Optimalidad, 378, 386, 394
Orden invertido de bits, 196
Ortogonal, 141-200
- base, 141
- complemento, 145-146
- DVS, 148
- matriz, 175
- proyección, 152-159
- valores característicos, 272
- vectores unitarios, 141
- vectores y subespacios, 141-151
  - *También véase* Proceso de Gram-Schmidt

Ortogonalización, 174, 182, 187, 331, 375, 477, 481
Ortonormal, 141-143, 148, 174-188
Oscilación, 234, 270, 274-275

P

$PA = LU$, 38, 39
Panqueque, 152
Paralelogramo, 4
Paréntesis, 6, 21-24, 34, 45-49, 134, 213, 332, 434, 445, 476
Patrones escalonados, 78
Permutación, 37-45, 202-203, 211-218
Permutación impar, 226-227
Permutación par, 44, 217, 230, 436, 453
Perpendicular. *Véase* Ortogonal
Perron-Frobenius, 261, 262
Perturbación, 62, 353, 357
Pivoteo completo, 63
Pivoteo parcial, 62, 352
Pivotes, 311
- fórmulas para los pivotes, 202
- positivos, 318
- prueba, 47-49
- variables, 80-81, 384

Pivotes diferentes de cero, 48
Planos, 4-5
Planos paralelos, 7, 8
Polinomio, 389, 478, 480, 481
Polinomio de Legendre, 182, 185
Polinomios característicos, 235
Polinomios por partes, 347-348
Póquer, 377, 412-414
Positiva semidefinida, 314, 321
Potencial, 339, 349, 478
Potencial en los nodos, 115
Potencias de matrices, 255
Preacondicionador, 368
Precio imaginario, 393, 396
Primer pivote, 12
Primera búsqueda de amplitud, 406
Primera búsqueda de profundidad, 406
Principio de incertidumbre, 250
Principio de incertidumbre de Heisenberg, 250
Principio de Rayleigh, 342
Principio maximin, 344, 409, 411
Principios mínimos, 339-345
Problema con valor inicial, 233
Problema de dieta, 380
Problema de dos puntos con valor en la frontera, 59
Problema de la ruta mínima, 404
Problema de transporte, 381, 406
Problema del matrimonio, 403-405
Problema dual, 382-391
Problema primal, 392
Proceso de Gram-Schmidt, 174-187
Proceso de Markov, 238, 257-259
Proceso de Markov continuo, 273
Producto. *Véase* multiplicación de matrices
Producto cartesiano, 417
Producto de Kronecker, 418
Producto interno, 20, 143, 169
Producto interno de funciones, 183
Producto punto. *Véase* Producto interno
Programación dinámica, 406
Programación lineal, 377-414
- desigualdades lineales, 377-381
- modelos de redes, 401-407
- método simplex, 382-291
- problema dual, 382-391
- restricciones, 378-380
- tabla (tableau), 386-388
- teoría de juegos, 408-413

Promedio ponderado, 169
Proyección, 322, 328, 338, 390-392, 416, 447-448, 450, 461, 465, 467, 475, 479, 481
Proyección sobre una recta, 152-159
Prueba de detención, 386, 388, 391, 471
Pruebas para la característica de ser positiva definida, 318-330
Punto de intersección, 4, 5
Punto mínimo, 311
Puntos silla, 311-317, 408

Q

Químico, 156, 203, 273
$Q\Lambda Q^T$, 320-323, 327

R

Radio espectral, 351
Raíces de la unidad, 189-190
Rango como espacio columna, 92
Rango de los renglones = rango de las columnas, 105
Rango de una matriz, 83, 98, 104
Rango total, 103, 109
Rango uno, 87, 107-114, 138, 140, 329, 333, 337, 417-418, 438-439, 464, 473, 479, 480
Razones de determinantes, 1, 222, 224
Red, 114-119, 124, 401-407, 478
Regla de convolución, 189
Regla de Cramer, 202, 221-222
Regla de las columnas, 21
Renglón a la vez, 372
Renglón multiplicado por una columna, 20
Representación por renglón, 428, 480
Representación por columna, 7, 8
Reescalamiento, 390
Restricción, 82, 85, 340-344, 346, 378-387, 390, 392, 394-402, 406, 470-471, 480
Restricciones de igualdad, 383
$\mathbf{R}^n$, 69, 72-73, 288
Rotación del plano, 361, 365, 367
RR^T y R^TR, 51, 52

S

$S^{-1}AS$, 132, 245-248, 285, 293, 299, 301, 324, 477
Semiancho de banda, 61
Semidefinida, 314, 321, 327, 329, 333-334, 467, 475, 480, 488
Semiespacio, 377
Serie de Fourier, 182
Serie de Fourier discreta, 192
Serie de Taylor, 315
Seudoinversa, 108, 148, 161, 335
Signos de los valores característicos, 271, 308, 311, 314, 318, 324-326, 329, 346, 478
Sistema sobredeterminado, 153, 166
Sistemas incurables, 13
Sobredeterminado, 153, 166
Sobrerrelajamiento, 368-371

Sobrerrelajamiento sucesivo (SRS), 368, 369
Soluciones especiales, 80, 81, 104
Soluciones particulares, 82, 83
Subdeterminado, 161
Subespacio, 70, 98
- fundamental, 102-115, 123, 137-139, 187, 392, 477, 481
- ortogonal, 114, 141-200, 399, 446, 477, 479, 481

Subespacio ortogonal, 143-149, 151-152, 415, 479
Subespacios fundamentales, 102-113
Submatriz, 44, 78, 148, 196, 213, 223, 224, 296, 318-319, 346, 404, 407, 433, 438, 450, 472
Submatriz principal, 87
Sucesión de Krylov, 365
Suma, 7-8, 21, 70-73, 82, 115, 126-127, 176-178, 181-184
Suma de cuadrados, 142, 160, 166, 177, 182, 199, 318-323, 327, 464
Suma de espacios, 415-421
Suma de vectores, 6
Sumatoria, 142, 160, 168, 177, 182, 199, 318-320, 327
Superposición, 237
Sustitución hacia atrás, 12, 36

T

Tabla (tableau), 386-388
Tablero de ajedrez, 139, 216, 242
Tensor, 2, 418
Teorema espectral, 285
Teorema flujo máx-corte mín, 402
Teorema fundamental del álgebra lineal, 106, 116-117, 141, 146-147, 335, 398
Teorema minimax, 344, 393, 409, 411
Teoría de juegos, 377-414
Teoría de von Neumann, 412
Tetraedro, 158, 471
TFR. *Véase* Transformada de Fourier rápida
Thorp, 412
Transformación, 126-129, 131-136
Transformación de semejanza, 293-306
Transformación identidad, 294
Transformación lineal, 125-137
Transformada de Fourier discreta, 188, 189, 287
Transformada de Fourier rápida (TFR), 188-197, 287, 372, 419, 475, 477
- identidad fundamental, 287
- ortogonalidad, 188-197

Traspuesta conjugada, 283
Traza, 239-241, 243-244, 250-253
Tukey, John, 194

U

Unicidad, 69

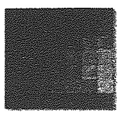
V

Valor característico cero, 109, 236, 238, 241-243, 246, 488
Valor del juego, 409
Valores característicos diferentes de cero, 49
Valores característicos distintos, 245-247, 297-299, 303, 308, 427, 473
Valores característicos dobles, 246, 422
Valores característicos repetidos, 246
Valores característicos y vectores característicos, 233-309
- A^k, 255-265
- cálculo, 359-366
- diagonalización, 245-254
- forma de Jordan, 300, 422
- inestabilidad, 234, 259, 270-273
- polinomio característico, 235
- matrices complejas, 280-292
- matriz positiva definida, 311
- prueba del valor característico, 320
- semejanza, 293-306
- valores característicos dobles, 246, 422

Variable de entrada, 387
Variable de salida, 387
Variable floja, 379, 383
Variable libre, 80-85, 89-91, 94, 104-109, 124, 384-387, 396, 437-441, 477-481
Varianzas y mínimos cuadrados ponderados, 169
Vector, 2-9, 29-24
Vector cero, 69
Vector de costo, 382
Vector error, 161-162, 165-167, 170-172
Vector unitario, 143, 174
Vectores característicos generalizados, 268
Vectores columna, 6-7, 20
Vértice de un conjunto factible, 381-391, 395-397
Volumen, 201

W

Wilkinson, 355
Wronskiano, 268



Álgebra lineal en pocas palabras

(*A es de n por n*)

No singular	**Singular**
A es invertible.	A no es invertible.
Las columnas son independientes.	Las columnas son dependientes.
Los renglones son independientes.	Los renglones son dependientes.
El determinante es diferente de cero.	El determinante es cero.
$Ax = 0$ tiene una solución $x = 0$.	$Ax = 0$ tiene una infinidad de soluciones.
$Ax = b$ tiene una solución $x = A^{-1}b$.	$Ax = b$ no tiene solución o tiene una infinidad de soluciones.
A tiene n pivotes (diferentes de cero).	A tiene $r < n$ pivotes.
A tiene rango completo $r = n$.	A tiene rango $r < n$.
La forma escalonada reducida por renglones es $R = I$.	R tiene por lo menos un renglón de ceros.
El espacio columna es todo $\mathbf{R}^n$.	La dimensión del espacio columna es $r < n$.
El espacio renglón es todo $\mathbf{R}^n$.	El espacio renglón tiene dimensión $r < n$.
Todos los valores característicos son diferentes de cero.	Cero es un valor característico de A.
$A^{\mathrm{T}}A$ es positiva definida simétrica.	$A^{\mathrm{T}}A$ sólo es semidefinida.
A tiene n valores singulares (positivos).	A tiene $r < n$ valores singulares.

Cada recta de la columna singular puede hacerse cuantitativa usando r.